GenCore version 5.1.6


OM protein - protein search, using sw model

Run on: October 2, 2003, 09:35:22 ; Search time 23 Seconds
(without alignments)
61.339 Million cell updates/sec

Title: US-09-939-293A-19_COPY_56_85
Perfect score: 141
Sequence: 1 AVPIAQKSEPHSLSSEALMRRAVSLVTDST 30

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | | Description |
|---|---|---|---|---|---|---|---|
| 1 | 141 | 100.0 | 239 | 1 | SMAC_HUMAN | Q9nr28 | homo sapien |
| 2 | 138 | 97.9 | 237 | 1 | SMAC_MOUSE | Q9jiq3 | mus musculu |
| 3 | 53 | 37.6 | 556 | 1 | YMC3_YEAST | Q03718 | saccharomyc |
| 4 | 51 | 36.2 | 455 | 1 | YNE2_CAEEL | P30641 | caenorhabdi |
| 5 | 50 | 35.5 | 429 | 1 | ELK1_MOUSE | P41969 | mus musculu |
| 6 | 47 | 33.3 | 131 | 1 | YI38_MYCTU | Q50595 | mycobacteri |
| 7 | 47 | 33.3 | 250 | 1 | LINC_PSEPA | P50197 | pseudomonas |
| 8 | 47 | 33.3 | 353 | 1 | CV04_MOUSE | Q8r5a6 | mus musculu |
| 9 | 47 | 33.3 | 944 | 1 | YMH6_YEAST | Q03631 | saccharomyc |
| 10 | 45.5 | 32.3 | 1593 | 1 | AT12_HUMAN | P58397 | homo sapien |
| 11 | 45 | 31.9 | 124 | 1 | VAF2_DROME | Q9vnl3 | drosophila |
| 12 | 45 | 31.9 | 318 | 1 | NRFD_ECOLI | P32709 | escherichia |
| 13 | 45 | 31.9 | 497 | 1 | CV04_MACFA | Q95ki1 | macaca fasc |
| 14 | 45 | 31.9 | 2515 | 1 | TUD_DROME | P25823 | drosophila |
| 15 | 45 | 31.9 | 3122 | 1 | DPOZ_MOUSE | Q61493 | mus musculu |
| 16 | 44 | 31.2 | 127 | 1 | VATF_ANOGA | Q17029 | anopheles g |
| 17 | 44 | 31.2 | 229 | 1 | CLC1_SCHPO | Q9usp6 | schizosacch |
| 18 | 44 | 31.2 | 440 | 1 | DNAA_THEMA | P46798 | thermotoga |
| 19 | 44 | 31.2 | 483 | 1 | KPYK_METEX | O05118 | methylobact |
| 20 | 44 | 31.2 | 522 | 1 | IBMP_CAMVJ | P13218 | cauliflower |
| 21 | 44 | 31.2 | 554 | 1 | UL25_HSV7J | P52388 | human herpe |
| 22 | 44 | 31.2 | 621 | 1 | DCTB_RHIME | P13633 | rhizobium m |
| 23 | 44 | 31.2 | 799 | 1 | YJV2_YEAST | P40897 | saccharomyc |
| 24 | 44 | 31.2 | 878 | 1 | SYL_TREPA | O83595 | treponema p |
| 25 | 44 | 31.2 | 1415 | 1 | RPOC_HAEIN | P43739 | haemophilus |
| 26 | 44 | 31.2 | 1518 | 1 | KKK1_YEAST | P34244 | saccharomyc |
| 27 | 43.5 | 30.9 | 336 | 1 | PYRD_ECOL6 | Q8fj91 | escherichia |
| 28 | 43.5 | 30.9 | 336 | 1 | PYRD_ECOLI | P05021 | escherichia |
| 29 | 43.5 | 30.9 | 522 | 1 | Z238_HUMAN | Q99592 | homo sapien |
| 30 | 43.5 | 30.9 | 522 | 1 | Z238_MOUSE | Q9wuk6 | mus musculu |
| 31 | 43.5 | 30.9 | 522 | 1 | Z238_RAT | Q9jky3 | rattus norv |
| 32 | 43.5 | 30.9 | 2298 | 1 | CU05_HUMAN | Q9y3r5 | homo sapien |
| 33 | 43 | 30.5 | 317 | 1 | TYSY_CRYNE | P45351 | cryptococcu |
| 34 | 43 | 30.5 | 331 | 1 | PLSX_VIBVU | Q8d8g5 | vibrio vuln |
| 35 | 43 | 30.5 | 334 | 1 | BC12_HUMAN | Q9hb09 | homo sapien |
| 36 | 43 | 30.5 | 424 | 1 | YL52_YERPE | Q8zel2 | yersinia pe |
| 37 | 43 | 30.5 | 434 | 1 | MOTC_RHIME | Q52963 | rhizobium m |
| 38 | 43 | 30.5 | 505 | 1 | VP5_AHSV4 | Q02168 | african hor |
| 39 | 43 | 30.5 | 658 | 1 | VATI_METTH | O27041 | methanobact |
| 40 | 43 | 30.5 | 1260 | 1 | ALS1_CANAL | P46590 | candida alb |
| 41 | 42.5 | 30.1 | 378 | 1 | PDXB_ECOLI | P05459 | escherichia |
| 42 | 42.5 | 30.1 | 608 | 1 | XYNC_FIBSU | P35811 | fibrobacter |
| 43 | 42.5 | 30.1 | 689 | 1 | SYM_HALN1 | Q9hsa4 | halobacteri |
| 44 | 42 | 29.8 | 110 | 1 | VATF_XENLA | Q9i8h3 | xenopus lae |
| 45 | 42 | 29.8 | 119 | 1 | VATF_RAT | P50408 | rattus norv |

ALIGNMENTS

RESULT 1
SMAC_HUMAN
ID SMAC_HUMAN STANDARD; PRT; 239 AA.
AC Q9NR28; Q96LV0; Q9BT11; Q9HAV6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Smac protein, mitochondrial precursor (Second mitochondria-derived
DE activator of caspase) (Direct IAP binding protein with low pI).
GN SMAC OR DIABLO.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1), PARTIAL SEQUENCE, FUNCTION, AND TISSUE
RP SPECIFICITY.
RX MEDLINE=20383536; PubMed=10929711;
RA Du C., Fang M., Li Y., Li L., Wang X.;
RT "Smac, a mitochondrial protein that promotes cytochrome c-dependent
RT caspase activation by eliminating IAP inhibition.";
RL Cell 102:33-42(2000).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RA Watanabe K., Kumagai A., Itakura S., Yamazaki M., Tashiro H., Ota T.,
RA Suzuki Y., Obayashi M., Nishi T., Shibahara T., Tanaka T.,
RA Nakamura Y., Isogai T., Sugano S.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 2), AND CHARACTERIZATION.
RX PubMed=10950947;
RA Srinivasula S.M., Datta P., Fan X.J., Fernandes-Alnemri T., Huang Z.,
RA Alnemri E.S.;
RT "Molecular determinants of the caspase-promoting activity of
RT Smac/DIABLO and its role in the death receptor pathway.";
RL J. Biol. Chem. 275:36152-36157(2000).
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Cerebellum;
RA Nishi T., Nakagawa S., Senoh A., Mizuguchi H., Inagaki H., Suzuki Y.,
RA Hata H., Nakagawa K., Mizuno S., Morinaga M., Kawamura M.,
RA Sugiyama T., Irie R., Otsuki T., Sato H., Nishikawa T., Sugiyama A.,
RA Kawakami B., Nagai K., Isogai T., Sugano S.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
RN [5]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Muscle, and Uterus;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

```
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [6]
RP   X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 56-239.
RX   MEDLINE=20426096; PubMed=10972280;
RA   Chai J., Du C., Wu J.W., Kyin S., Wang X., Shi Y.;
RT   "Structural and biochemical basis of apoptotic activation by
RT   Smac/DIABLO.";
RL   Nature 406:855-862(2000).
RN   [7]
RP   STRUCTURE BY NMR OF 56-64 IN COMPLEX WITH BIRC4.
RX   MEDLINE=21020961; PubMed=11140637;
RA   Liu Z., Sun C., Olejniczak E.T., Meadows R.P., Betz S.F., Oost T.,
RA   Herrmann J., Wu J.C., Fesik S.W.;
RT   "Structural basis for binding of Smac/DIABLO to the XIAP BIR3
RT   domain.";
RL   Nature 408:1004-1008(2000).
CC   -!- FUNCTION: PROMOTES APOPTOSIS BY ACTIVATING CASPASES IN THE
CC       CYTOCHROME C/APAF-1/CASPASE-9 PATHWAY. ACTS BY OPPOSING THE
CC       INHIBITORY ACTIVITY OF INHIBITOR OF APOPTOSIS PROTEINS (IAP).
CC   -!- SUBUNIT: Homodimer. Interacts with BIRC2, BIRC3, BIRC4/XIAP and
CC       BIRC7.
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL BUT RELEASED INTO THE CYTOSOL
CC       WHEN CELLS UNDERGO APOPTOSIS.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=1;
CC         IsoId=Q9NR28-1; Sequence=Displayed;
CC       Name=2; Synonyms=Diablo-S;
CC         IsoId=Q9NR28-2; Sequence=VSP_004397;
CC   -!- TISSUE SPECIFICITY: UBIQUITOUSLY EXPRESSED WITH HIGHEST EXPRESSION
CC       IN TESTIS. EXPRESSION IS ALSO HIGH IN HEART, LIVER, KIDNEY,
CC       SPLEEN, PROSTATE AND OVARY. LOW IN BRAIN, LUNG, THYMUS AND
CC       PERIPHERAL BLOOD LEUKOCYTES.
CC   -!- DOMAIN: The mature N-terminus mediates interaction with
CC       BIRC4/XIAP.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF262240; AAF87716.1; -.
DR   EMBL; AK024768; BAB14994.1; -.
DR   EMBL; AF298770; AAG22077.1; -.
DR   EMBL; AK057778; BAB71568.1; -.
DR   EMBL; BC004417; AAH04417.1; -.
DR   PDB; 1FEW; 13-SEP-00.
DR   PDB; 1G3F; 10-JAN-01.
DR   PDB; 1G73; 10-JAN-01.
DR   MIM; 605219; -.
DR   GO; GO:0005739; C:mitochondrion; TAS.
DR   GO; GO:0008635; P:caspase activation via cytochrome c; TAS.
DR   GO; GO:0008625; P:induction of apoptosis via death domain rec. . .; TAS.
DR   GO; GO:0006917; P:induction of apoptosis; TAS.
KW   Transit peptide; Mitochondrion; Apoptosis; Alternative splicing;
KW   3D-structure.
FT   TRANSIT       1     55       MITOCHONDRION.
FT   CHAIN        56    239       SMAC PROTEIN.
FT   SITE         56     60       IAP-BINDING MOTIF (BY SIMILARITY).
FT   VARSPLIC      1     60       MAALKSWLSRSVTSFFRYRQCLCVPVVANFKKRCFSELIRP
FT                                WHKTVTIGFGVTLCAVPIA -> MKSDFYF (in
FT                                isoform 2).
FT                                /FTId=VSP_004397.
FT   CONFLICT     32     32       K -> E (IN REF. 4).
FT   CONFLICT     44     44       K -> R (IN REF. 2).
FT   CONFLICT     62    105       MISSING (IN REF. 4).
FT   CONFLICT    165    165       E -> K (IN REF. 4).
SQ   SEQUENCE   239 AA;  27131 MW;  70C2AE0DC654D031 CRC64;

  Query Match             100.0%;  Score 141;  DB 1;  Length 239;
  Best Local Similarity   100.0%;  Pred. No. 3e-14;
  Matches   30;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AVPIAQKSEPHSLSSEALMRRAVSLVTDST 30
              |||||||||||||||||||||||||||||
Db         56 AVPIAQKSEPHSLSSEALMRRAVSLVTDST 85


RESULT 2
SMAC_MOUSE
ID   SMAC_MOUSE     STANDARD;      PRT;   237 AA.
AC   Q9JIQ3; Q9CZD1; Q9DCD3;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Smac protein, mitochondrial precursor (Second mitochondria-derived
DE   activator of caspase) (Direct IAP binding protein with low pI).
GN   SMAC OR DIABLO.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A., FUNCTION, SUBCELLULAR LOCATION, AND TISSUE
RP   SPECIFICITY.
RC   STRAIN=BALB/c; TISSUE=Kidney;
RX   MEDLINE=20383537; PubMed=10929712;
RA   Verhagen A.M., Ekert P.G., Pakusch M., Silke J., Connolly L.M.,
RA   Reid G.E., Moritz R.L., Simpson R.J., Vaux D.L.;
RT   "Identification of DIABLO, a mammalian protein that promotes apoptosis
RT   by binding to and antagonizing IAP proteins.";
RL   Cell 102:43-53(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J;
RX   MEDLINE=21085660; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA   Suzuki H., Toyo-Oka K., Wang K.H., Weitz C., Whittaker C.,
RA   Wilming L., Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H.,
RA   Kohtsuki S.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
CC   -!- FUNCTION: PROMOTES APOPTOSIS BY ACTIVATING CASPASES IN THE
CC       CYTOCHROME C/APAF-1/CASPASE-9 PATHWAY. ACTS BY OPPOSING THE
CC       INHIBITORY ACTIVITY OF INHIBITOR OF APOPTOSIS PROTEINS (IAP).
CC   -!- SUBUNIT: Homodimer. Interacts with BIRC2, BIRC3, BIRC4/XIAP and
```

```
CC       BIRC7 (By similarity).
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL BUT RELEASED INTO THE CYTOSOL
CC       WHEN CELLS UNDERGO APOPTOSIS.
CC   -!- TISSUE SPECIFICITY: HIGHEST EXPRESSION FOUND IN HEART, LIVER,
CC       KIDNEY AND TESTIS.
CC   -!- DOMAIN: The mature N-terminus mediates interaction with
CC       BIRC4/XIAP (By similarity).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF203914; AAF82190.1; -.
DR   EMBL; AK012760; BAB28450.1; -.
DR   EMBL; AK002887; -; NOT_ANNOTATED_CDS.
DR   HSSP; Q9NR28; 1FEW.
DR   MGD; MGI:1913843; 0610041G12Rik.
KW   Transit peptide; Mitochondrion; Apoptosis.
FT   TRANSIT       1     53       MITOCHONDRION (BY SIMILARITY).
FT   CHAIN        54    237       SMAC PROTEIN.
FT   SITE         54     58       IAP-BINDING MOTIF (BY SIMILARITY).
FT   CONFLICT     64     64       H -> Q (IN REF. 2).
SQ   SEQUENCE   237 AA;  26829 MW;  E53E6F04F1C390A1 CRC64;

  Query Match             97.9%;  Score 138;  DB 1;  Length 237;
  Best Local Similarity   96.7%;  Pred. No. 8.5e-14;
  Matches   29;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      .       1 AVPIAQKSEPHSLSSEALMRRAVSLVTDST 30
                  |||||||||||||||:|||||||||||||||
Db          54 AVPIAQKSEPHSLSNEALMRRAVSLVTDST 83


RESULT 3
YMC3_YEAST
ID   YMC3_YEAST     STANDARD;      PRT;   556 AA.
AC   Q03718;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Hypothetical 64.0 kDa protein in RPS17A-APT1 intergenic region.
GN   YML023C.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288c / AB972;
RX   PubMed=9169872;
RA   Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T.,
RA   Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S.,
RA   Jagels K., Lye G., Moule S., Odell C., Pearson D., Rajandream M.A.,
RA   Rice P., Skelton J., Walsh S., Whitehead S., Barrell B.G.;
RT   "The nucleotide sequence of Saccharomyces cerevisiae chromosome
RT   XIII.";
RL   Nature 387:90-93(1997).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z46659; CAA86632.1; -.
DR   PIR; S49754; S49754.
DR   SGD; S0004485; YML023C.
KW   Hypothetical protein; Transmembrane.
FT   TRANSMEM     61     81       POTENTIAL.
FT   TRANSMEM    483    503       POTENTIAL.
SQ   SEQUENCE   556 AA;  64012 MW;  6680A04D91E537DD CRC64;

  Query Match             37.6%;  Score 53;  DB 1;  Length 556;
  Best Local Similarity   41.7%;  Pred. No. 2.5;
  Matches   10;  Conservative    7;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          3 PIAQKSEPHSLSSEALMRRAVSLV 26
              || : ::|:: | | | |||: |:
Db         85 PILRANDPYNTSRETLSRRALKLL 108


RESULT 4
YNE2_CAEEL
ID   YNE2_CAEEL     STANDARD;      PRT;   455 AA.
AC   P30641;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-APR-1993 (Rel. 25, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Hypothetical protein R08D7.2 in chromosome III.
GN   R08D7.2.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RX   MEDLINE=94150718; PubMed=7906398;
RA   Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA   Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA   Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
RA   Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA   Johnston L., Jones M., Kershaw J., Kirsten J., Laisster N.,
RA   Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA   Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
RA   Sims M., Smaldon N., Smith A., Smith M., Sonnhammer E., Staden R.,
RA   Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA   Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,
RA   Wohldman P.;
RT   "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT   elegans.";
RL   Nature 368:32-38(1994).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z12017; CAA78048.1; -.
DR   PIR; S41037; S24458.
DR   WormPep; R08D7.2; CE00290.
DR   Pfam; PF04181; DUF408; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   455 AA;  52438 MW;  1FDFADAA58980F3E CRC64;

  Query Match             36.2%;  Score 51;  DB 1;  Length 455;
  Best Local Similarity   48.1%;  Pred. No. 4;
  Matches   13;  Conservative    5;  Mismatches    7;  Indels    2;  Gaps    1;

Qy          3 PIAQKSEPHSLSSEALMRRAVSLVTDS 29
              |||:|||| : |:     :: |||||
Db        141 PIAEKSEP--VDSQFKKEPSIWLVTDS 165


RESULT 5
ELK1_MOUSE
ID   ELK1_MOUSE     STANDARD;      PRT;   429 AA.
```

```
AC   P41969;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   ETS-domain protein ELK-1.
GN   ELK1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6; TISSUE=Embryo;
RX   MEDLINE=97017146; PubMed=8863747;
RA   Grevin D., Ung S., Denhez F., Dehem M., Quatannens B., Begue A.,
RA   Stehelin D., Martin P.;
RT   "Structure and organization of the mouse elk1 gene.";
RL   Gene 174:185-188(1996).
RN   [2]
RP   SEQUENCE OF 5-224 FROM N.A.
RC   TISSUE=Embryo;
RX   MEDLINE=95047310; PubMed=7958835;
RA   Giovane A., Pintzas A., Maira S.-M., Sobieszczuk P., Wasylyk B.;
RT   "Net, a new ets transcription factor that is activated by Ras.";
RL   Genes Dev. 8:1502-1513(1994).
CC   -!- FUNCTION: STIMULATES TRANSCRIPTION. BINDS TO PURINE-RICH DNA
CC       SEQUENCES. CAN FORM A TERNARY COMPLEX WITH THE SERUM RESPONSE
CC       FACTOR AND THE ETS AND SRF MOTIFS OF THE FOS SERUM RESPONSE
CC       ELEMENT.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN THE BRAIN, AND TO A
CC       LESSER EXTENT IN THE HEART, LIVER AND MUSCLE.
CC   -!- SIMILARITY: BELONGS TO THE ETS FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X87257; CAA60715.1; -.
DR   EMBL; Z36939; CAA85391.1; -.
DR   PIR; JC4965; JC4965.
DR   HSSP; P28324; 1BC8.
DR   TRANSFAC; T05013; -.
DR   MGD; MGI:101833; Elk1.
DR   InterPro; IPR000418; Ets.
DR   InterPro; IPR002341; HSF_ETS.
DR   Pfam; PF00178; Ets; 1.
DR   PRINTS; PR00454; ETSDOMAIN.
DR   SMART; SM00413; ETS; 1.
DR   PROSITE; PS00345; ETS_DOMAIN_1; 1.
DR   PROSITE; PS00346; ETS_DOMAIN_2; 1.
DR   PROSITE; PS50061; ETS_DOMAIN_3; 1.
KW   Transcription regulation; Activator; Nuclear protein; DNA-binding;
KW   Phosphorylation.
FT   DNA_BIND      5     86       ETS-DOMAIN.
FT   CONFLICT    133    133       P -> T (IN REF. 2).
SQ   SEQUENCE   429 AA;  45243 MW;  B61B5B977731D54F CRC64;

  Query Match             35.5%;  Score 50;  DB 1;  Length 429;
  Best Local Similarity   38.5%;  Pred. No. 5.3;
  Matches   10;  Conservative    8;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          5 AQKSEPHSLSSEALMRRAVSLVTDST 30
              | |:||   :|| |: |  :::|::|
Db        270 AVKAEPEVSASEGLLARLPAILTENT 295


RESULT 6
YI38_MYCTU
ID   YI38_MYCTU     STANDARD;      PRT;   131 AA.
AC   Q50595; P95168;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Hypothetical protein Rv1838c.
GN   RV1838C OR MT1886 OR MTCY1A11.05 OR MTCY359.35.
OS   Mycobacterium tuberculosis.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX   NCBI_TaxID=1773;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=H37Rv;
RX   MEDLINE=98295987; PubMed=9634230;
RA   Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA   Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekaia F.,
RA   Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA   Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA   Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA   Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA   Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA   Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT   "Deciphering the biology of Mycobacterium tuberculosis from the
RT   complete genome sequence.";
RL   Nature 393:537-544(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CDC 1551 / Oshkosh;
RA   Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA   Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA   Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA   Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA   Bishai W.;
RT   "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT   laboratory strains.";
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z83859; CAB06116.1; -.
DR   EMBL; AE007047; AAK46157.1; -.
DR   PIR; F70663; F70663.
DR   TIGR; MT1886; -.
DR   TubercuList; Rv1838c; -.
DR   InterPro; IPR002716; PIN.
DR   InterPro; IPR006596; PINc.
DR   Pfam; PF01850; PIN; 1.
DR   SMART; SM00670; PINc; 1.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   131 AA;  14726 MW;  C164346E951BFE7E CRC64;

  Query Match             33.3%;  Score 47;  DB 1;  Length 131;
  Best Local Similarity   42.3%;  Pred. No. 3.8;
  Matches   11;  Conservative    5;  Mismatches    6;  Indels    4;  Gaps    1;

Qy          8 SEPHSLSSEALMRRAVS----LVTDS 29
              | || | :: |:  |:|    ||||:
Db         16 SHPHKLDAQRLLESALSGGERLVTDA 41


RESULT 7
LINC_PSEPA
ID   LINC_PSEPA     STANDARD;      PRT;   250 AA.
AC   P50197;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
```

```
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   2,5-dichloro-2,5-cyclohexadiene-1,4-diol dehydrogenase (EC 1.1.-.-)
DE   (2,5-DDOL dehydrogenase).
GN   LINC.
OS   Pseudomonas paucimobilis (Sphingomonas paucimobilis).
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Sphingomonadales;
OC   Sphingomonadaceae; Sphingomonas.
OX   NCBI_TaxID=13689;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=UT26;
RX   MEDLINE=94252977; PubMed=7515041;
RA   Nagata Y., Ohtomo R., Miyauchi K., Fukuda M., Yano K., Takagi M.;
RT   "Cloning and sequencing of a 2,5-dichloro-2,5-cyclohexadiene-1,4-diol
RT   dehydrogenase gene involved in the degradation of gamma-
RT   hexachlorocyclohexane in Pseudomonas paucimobilis.";
RL   J. Bacteriol. 176:3117-3125(1994).
CC   -!- FUNCTION: DEGRADATION OF 2,5-DICHLORO-2,5-CYCLOHEXADIENE-1,4-DIOL
CC       (2,5-DDOL) INTO 2,5-DICHLOROHYDROQUINONE (2,5-DCHQ).
CC   -!- PATHWAY: Gamma-hexachlorocyclohexane degradation; third step.
CC   -!- SIMILARITY: Belongs to the short-chain dehydrogenases/reductases
CC       (SDR) family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D14595; BAA03444.1; -.
DR   HSSP; P19992; 1HDC.
DR   InterPro; IPR002198; ADH_short.
DR   Pfam; PF00106; adh_short; 1.
DR   PRINTS; PR00080; SDRFAMILY.
DR   PROSITE; PS00061; ADH_SHORT; 1.
KW   Aromatic hydrocarbons catabolism; Oxidoreductase; NAD.
FT   NP_BIND      9     34       NAD (BY SIMILARITY).
FT   ACT_SITE   154    154       BY SIMILARITY.
SQ   SEQUENCE   250 AA;  25644 MW;  FFC1CAEB47DF789D CRC64;

  Query Match             33.3%;  Score 47;  DB 1;  Length 250;
  Best Local Similarity   44.8%;  Pred. No. 8.1;
  Matches   13;  Conservative    3;  Mismatches   11;  Indels    2;  Gaps    1;

Qy          3 PIAQKSEPHSLSSEA--LMRRAVSLVTDS 29
              || : ||||  :  |  |:  | | || |
Db        211 PIGRFSEPHEQAQAAVWLLSDAASFVTGS 239


RESULT 8
CV04_MOUSE
ID   CV04_MOUSE     STANDARD;      PRT;   353 AA.
AC   Q8R5A6;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   TBC1 domain family protein C22orf4 homolog (Fragment).
GN   C22ORF4.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Mammary gland;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- SIMILARITY: Contains 1 Rab-GAP TBC domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; BC023106; AAH23106.1; ALT_INIT.
DR   InterPro; IPR000195; RabGAP_TBC.
DR   Pfam; PF00566; TBC; 1.
DR   SMART; SM00164; TBC; 1.
DR   PROSITE; PS50086; TBC_RABGAP; 1.
FT   NON_TER       1       1
FT   DOMAIN       58     282       RAB-GAP TBC.
SQ   SEQUENCE   353 AA;  41478 MW;  B16BD293761D4A53 CRC64;

  Query Match             33.3%;  Score 47;  DB 1;  Length 353;
  Best Local Similarity   39.1%;  Pred. No. 12;
  Matches    9;  Conservative    6;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          2 VPIAQKSEPHSLSSEALMRRAVS 24
              | ::  |:||:|:  || :|  |
Db         10 VTLSGTSDPHALADSALSKRETS 32


RESULT 9
YMH6_YEAST
ID   YMH6_YEAST     STANDARD;      PRT;   944 AA.
AC   Q03631;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Putative 107.6 kDa transcriptional regulatory protein in CPR3-HMG1
DE   intergenic region.
GN   YML076C.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288c / AB972;
RX   PubMed=9169872;
RA   Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T.,
RA   Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S.,
RA   Jagels K., Lye G., Moule S., Odell C., Pearson D., Rajandream M.A.,
RA   Rice P., Skelton J., Walsh S., Whitehead S., Barrell B.G.;
RT   "The nucleotide sequence of Saccharomyces cerevisiae chromosome
RT   XIII.";
RL   Nature 387:90-93(1997).
CC   -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC   -!- SIMILARITY: Contains 1 Zn(2)-Cys(6) fungal-type binuclear cluster
CC       domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
```

```
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; Z46373; CAA86502.1; -.
DR   PIR; S48821; S48821.
DR   SGD; S0004541; YML076C.
DR   InterPro; IPR001138; Fungi_TrN.
DR   SMART; SM00066; GAL4; 1.
DR   PROSITE; PS00463; ZN2_CY6_FUNGAL_1; 1.
DR   PROSITE; PS50048; ZN2_CY6_FUNGAL_2; FALSE_NEG.
KW   Hypothetical protein; Transcription regulation; DNA-binding; Zinc;
KW   Nuclear protein; Metal-binding.
FT   DNA_BIND     76    109       ZN(2)-CYS(6), FUNGAL-TYPE.
SQ   SEQUENCE   944 AA;  107560 MW;  020A56745DF52CCC CRC64;

  Query Match             33.3%;  Score 47;  DB 1;  Length 944;
  Best Local Similarity   38.5%;  Pred. No. 39;
  Matches   10;  Conservative    5;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          3 PIAQKSEPHSLSSEALMRRAVSLVTD 28
              |:   | | |::|||| :    :| |
Db        252 PVPISSAPTSINSEALFKHRPKIVGD 277


RESULT 10
AT12_HUMAN
ID   AT12_HUMAN     STANDARD;      PRT;  1593 AA.
AC   P58397;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   ADAMTS-12 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase
DE   with thrombospondin motifs 12) (ADAM-TS 12) (ADAM-TS12).
GN   ADAMTS12.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Fetal lung;
RX   MEDLINE=21264577; PubMed=11279086;
RA   Cal S., Argueelles J.M., Fernandez P.L., Lopez-Otin C.;
RT   "Identification, characterization, and intracellular processing of
RT   ADAM-TS12, a novel human disintegrin with a complex structural
RT   organization involving multiple thrombospondin-1 repeats.";
RL   J. Biol. Chem. 276:17932-17940(2001).
CC   -!- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
CC       matrix (By similarity).
CC   -!- TISSUE SPECIFICITY: Expressed exclusively in fetal lung. Is widely
CC       expressed in gastric carcinomas and in cancer cells of diverse
CC       origin.
CC   -!- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE 1 DOMAINS ARE IMPORTANT
CC       FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX (BY
CC       SIMILARITY).
CC   -!- PTM: THE PRECURSOR IS CLEAVED BY A FURIN ENDOPEPTIDASE.
CC   -!- PTM: IS SUBJECTED TO AN INTRACELLULAR MATURATION PROCESS LEADIND
CC       TO A FRAGMENT CONTAINING THE N-TERMINAL REGION INCLUDING THE
CC       METALLOPROTEINASE, DISINTEGRIN-LIKE, CYS-RICH AND TS-1 DOMAINS AND
CC       THE C-TERMINAL FRAGMENT CONTAINING THE SPACER 2 AND THE FOUR TS-1
CC       DOMAINS.
CC   -!- SIMILARITY: Belongs to peptidase family M12B.
CC   -!- SIMILARITY: Contains 1 disintegrin-like domain.
CC   -!- SIMILARITY: Contains 1 PLAC domain.
CC   -!- SIMILARITY: Contains 8 TSP type-1 domains.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AJ250725; CAC20419.1; -.
DR   Genew; HGNC:14605; ADAMTS12.
DR   MIM; 606184; -.
DR   InterPro; IPR001762; Disintegrin.
DR   InterPro; IPR002870; Pep_M12B_propep.
DR   InterPro; IPR001590; Reprolysin.
DR   InterPro; IPR000884; TSP1.
DR   InterPro; IPR006025; Zn_MTpeptdse.
DR   Pfam; PF01562; Pep_M12B_propep; 1.
DR   Pfam; PF01421; Reprolysin; 1.
DR   Pfam; PF00090; tsp_1; 6.
DR   SMART; SM00209; TSP1; 8.
DR   PROSITE; PS50215; ADAM_MEPRO; 1.
DR   PROSITE; PS00546; CYSTEINE_SWITCH; FALSE_NEG.
DR   PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
DR   PROSITE; PS50214; DISINTEGRIN_2; FALSE_NEG.
DR   PROSITE; PS50092; TSP1; 6.
DR   PROSITE; PS00142; ZINC_PROTEASE; 1.
KW   Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
KW   Repeat; Extracellular matrix.
FT   SIGNAL       1     25       POTENTIAL.
FT   PROPEP      26    240       BY SIMILARITY.
FT   CHAIN      241   1593       ADAMTS-12.
FT   DOMAIN     241    464       METALLOPROTEASE.
FT   DOMAIN     465    544       DISINTEGRIN-LIKE.
FT   DOMAIN     542    597       TSP TYPE-1 1.
FT   DOMAIN     597    700       CYS-RICH.
FT   DOMAIN     701    826       SPACER 1.
FT   DOMAIN     823    882       TSP TYPE-1 2.
FT   DOMAIN     886    942       TSP TYPE-1 3.
FT   DOMAIN     943    996       TSP TYPE-1 4.
FT   DOMAIN     996   1315       SPACER 2.
FT   DOMAIN    1312   1365       TSP TYPE-1 5.
FT   DOMAIN    1367   1421       TSP TYPE-1 6.
FT   DOMAIN    1422   1470       TSP TYPE-1 7.
FT   DOMAIN    1471   1531       TSP TYPE-1 8.
FT   DOMAIN     302    305       POLY-GLU.
FT   DOMAIN    1538   1570       PLAC.
FT   SITE       208    208       CYSTEINE SWITCH (POTENTIAL).
FT   METAL      392    392       ZINC (CATALYTIC) (BY SIMILARITY).
FT   ACT_SITE   393    393       BY SIMILARITY.
FT   METAL      396    396       ZINC (CATALYTIC) (BY SIMILARITY).
FT   METAL      402    402       ZINC (CATALYTIC) (BY SIMILARITY).
FT   CARBOHYD   105    105       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   125    125       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   215    215       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   485    485       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   685    685       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   790    790       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   951    951       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD  1104   1104       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD  1275   1275       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD  1300   1300       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD  1320   1320       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD  1371   1371       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD  1378   1378       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD  1503   1503       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE  1593 AA;  177545 MW;  07F9F48E63BD83A3 CRC64;

  Query Match             32.3%;  Score 45.5;  DB 1;  Length 1593;
  Best Local Similarity   38.7%;  Pred. No. 1.2e+02;
  Matches   12;  Conservative    6;  Mismatches    6;  Indels    7;  Gaps    1;

Qy          1 AVPIAQKS-------EPHSLSSEALMRRAVS 24
                :| |:||       | |:| | :| ||::|
Db        213 SVNISQKQELWREKWERHNLPSRSLSRRSIS 243
```

```
RESULT 11
VAF2_DROME
ID   VAF2_DROME     STANDARD;      PRT;   124 AA.
AC   Q9VNL3;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Probable vacuolar ATP synthase subunit F 2 (EC 3.6.3.14) (V-ATPase F
DE   subunit 2) (Vacuolar proton pump F subunit 2) (V-ATPase 14 kDa subunit
DE   2).
GN   CG1076.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
CC   -!- FUNCTION: SUBUNIT OF THE PERIPHERAL V1 COMPLEX OF VACUOLAR ATPASE
CC       ESSENTIAL FOR ASSEMBLY OR CATALYTIC FUNCTION. V-ATPASE IS
CC       RESPONSIBLE FOR ACIDIFYING A VARIETY OF INTRACELLULAR COMPARTMENTS
CC       IN EUKARYOTIC CELLS (BY SIMILARITY).
CC   -!- CATALYTIC ACTIVITY: ATP + H(2)O + H(+)(In) = ADP + phosphate +
CC       H(+)(Out).
CC   -!- SUBUNIT: V-ATPase is an heteromultimeric enzyme composed of a
CC       peripheral catalytic V1 complex (components A to H) attached to
CC       an integral membrane V0 proton pore complex (components: a, c, c',
CC       c'' and d) (By similarity).
CC   -!- SIMILARITY: Belongs to the V-ATPase F subunit family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE003600; AAF51917.1; -.
DR   FlyBase; FBgn0037402; CG1076.
DR   InterPro; IPR005772; ATPsynt_F_euk.
DR   InterPro; IPR002841; ATPsynt_Fsub.
DR   Pfam; PF01990; ATP-synt_F; 1.
DR   ProDom; PD003811; ATPsynt_F_euk; 1.
DR   TIGRFAMs; TIGR01101; V_ATP_synt_F; 1.
KW   Hydrolase; ATP synthesis; Hydrogen ion transport.
SQ   SEQUENCE   124 AA;  14132 MW;  0C5093AFDF006BAB CRC64;

  Query Match             31.9%;  Score 45;  DB 1;  Length 124;
  Best Local Similarity   32.3%;  Pred. No. 7.2;
  Matches   10;  Conservative    8;  Mismatches    9;  Indels    4;  Gaps    1;

Qy          1 AVP----IAQKSEPHSLSSEALMRRAVSLVT 27
              |||    |  |  |:  | :::::||  ::|
Db         88 AVPTVLEIPSKQHPYDSSRDSILKRAQRVIT 118


RESULT 12
NRFD_ECOLI
ID   NRFD_ECOLI     STANDARD;      PRT;   318 AA.
AC   P32709;
DT   01-OCT-1993 (Rel. 27, Created)
DT   01-OCT-1993 (Rel. 27, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   NrfD protein.
GN   NRFD OR B4073.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94335626; PubMed=8057835;
RA   Hussain H.A., Grove J., Griffiths L., Busby S., Cole J.;
RT   "A seven-gene operon essential for formate-dependent nitrite
RT   reduction to ammonia by enteric bacteria.";
RL   Mol. Microbiol. 12:153-163(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / MG1655;
RX   MEDLINE=94089392; PubMed=8265357;
RA   Blattner F.R., Burland V.D., Plunkett G. III, Sofia H.J.,
RA   Daniels D.L.;
RT   "Analysis of the Escherichia coli genome. IV. DNA sequence of the
RT   region from 89.2 to 92.8 minutes.";
RL   Nucleic Acids Res. 21:5408-5417(1993).
CC   -!- FUNCTION: PROBABLY INVOLVED IN THE TRANSFER OF ELECTRONS FROM THE
CC       QUINONE POOL TO THE TYPE-C CYTOCHROMES.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC       (Potential).
CC   -!- SIMILARITY: TO W.SUCCINOGENES POLYSULFIDE REDUCTASE CHAIN C.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X72298; CAA51044.1; -.
DR   EMBL; U00006; AAC43167.1; -.
DR   EMBL; AE000480; AAC77043.1; -.
DR   PIR; H65215; D57987.
DR   EcoGene; EG11947; nrfD.
```

```
DR   InterPro; IPR005614; NrfD.
DR   Pfam; PF03916; NrfD; 1.
KW   Transmembrane; Inner membrane; Complete proteome.
FT   TRANSMEM      18     38       POTENTIAL.
FT   TRANSMEM      57     73       POTENTIAL.
FT   TRANSMEM      90    112       POTENTIAL.
FT   TRANSMEM     150    170       POTENTIAL.
FT   TRANSMEM     180    199       POTENTIAL.
FT   TRANSMEM     222    242       POTENTIAL.
FT   TRANSMEM     258    278       POTENTIAL.
FT   TRANSMEM     288    310       POTENTIAL.
FT   CONFLICT     141    141       V -> L (IN REF. 2).
FT   CONFLICT     202    202       R -> A (IN REF. 2).
SQ   SEQUENCE   318 AA;  35113 MW;  BC5B3EF031D5CE29 CRC64;

  Query Match             31.9%;  Score 45;  DB 1;  Length 318;
  Best Local Similarity   34.6%;  Pred. No. 22;
  Matches    9;  Conservative    7;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          1 AVPIAQKSEPHSLSSEALMRRAVSLV 26
              |: | |:| |||  :: : |  : :|
Db        200 AMRIRQRSNPHSTEAQFVHRMEIPVV 225


RESULT 13
CV04_MACFA
ID   CV04_MACFA     STANDARD;      PRT;   497 AA.
AC   Q95KI1;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   TBC1 domain family protein C22orf4 homolog (QtrA-11492) (Fragment).
GN   C22ORF4.
OS   Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Macaca.
OX   NCBI_TaxID=9541;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Temporal cortex;
RA   Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
RA   Suzuki Y., Sugano S., Hashimoto K.;
RT   "Isolation of full-length cDNA clones from macaque brain cDNA
RT   libraries.";
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Contains 1 Rab-GAP TBC domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AB060857; BAB46876.1; ALT_INIT.
DR   InterPro; IPR000195; RabGAP_TBC.
DR   Pfam; PF00566; TBC; 1.
DR   SMART; SM00164; TBC; 1.
DR   PROSITE; PS50086; TBC_RABGAP; 1.
FT   NON_TER       1      1
FT   DOMAIN      202    426       RAB-GAP TBC.
SQ   SEQUENCE   497 AA;  56810 MW;  47EF1098A998937A CRC64;

  Query Match             31.9%;  Score 45;  DB 1;  Length 497;
  Best Local Similarity   45.8%;  Pred. No. 37;
  Matches   11;  Conservative    3;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          1 AVPIAQKSEPHSLSSEALMRRAVS 24
              || :   |:| :||| ||  |  |
Db        153 AVTLGGTSDPGTLSSSALSEREAS 176


RESULT 14
TUD_DROME
ID   TUD_DROME      STANDARD;      PRT;  2515 AA.
AC   P25823;
DT   01-MAY-1992 (Rel. 22, Created)
DT   01-MAY-1992 (Rel. 22, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Maternal tudor protein.
GN   TUD.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92038995; PubMed=1936993;
RA   Golumbeski G.S., Bardsley A., Tax F., Boswell R.E.;
RT   "Tudor, a posterior-group gene of Drosophila melanogaster, encodes a
RT   novel protein and an mRNA localized during mid-oogenesis.";
RL   Genes Dev. 5:2060-2070(1991).
CC   -!- FUNCTION: REQUIRED DURING OOGENESIS FOR THE FORMATION OF
CC       PRIMORDIAL GERM CELLS AND FOR NORMAL ABDOMINAL SEGMENTATION.
CC   -!- DEVELOPMENTAL STAGE: EXPRESSED THROUGHOUT THE LIFE CYCLE.
CC   -!- MISCELLANEOUS: THE TUD MRNA ACCUMULATES WITHIN THE POSTERIOR
CC       REGION OF THE DEVELOPING OOCYTE DURING THE EARLY TO MIDDLE STAGES
CC       OF OOGENESIS.
CC   -!- SIMILARITY: Contains 9 Tudor domains.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X62420; CAA44286.1; -.
DR   PIR; A41519; A41519.
DR   HSSP; Q16637; 1G5V.
DR   FlyBase; FBgn0003891; tud.
DR   GO; GO:0019090; P:mitochondrial rRNA, mitochondrial export; IMP.
DR   GO; GO:0007315; P:pole plasm assembly; IMP.
DR   InterPro; IPR001097; Maternal_tudor.
DR   InterPro; IPR002999; Tudor.
DR   Pfam; PF00567; TUDOR; 10.
DR   SMART; SM00333; TUDOR; 10.
DR   PROSITE; PS50304; TUDOR; 9.
KW   Developmental protein; Repeat.
FT   DOMAIN      455    513       TUDOR 1.
FT   DOMAIN      641    696       TUDOR 2.
FT   DOMAIN     1062   1122       TUDOR 3.
FT   DOMAIN     1355   1414       TUDOR 4.
FT   DOMAIN     1662   1718       TUDOR 5.
FT   DOMAIN     1839   1898       TUDOR 6.
FT   DOMAIN     2023   2082       TUDOR 7.
FT   DOMAIN     2211   2269       TUDOR 8.
FT   DOMAIN     2392   2451       TUDOR 9.
SQ   SEQUENCE   2515 AA;  285236 MW;  683C100AD308BADA CRC64;

  Query Match             31.9%;  Score 45;  DB 1;  Length 2515;
  Best Local Similarity   30.8%;  Pred. No. 2.5e+02;
  Matches    8;  Conservative    7;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          2 VPIAQKSEPHSLSSEALMRRAVSLVT 27
              :|: |: ||   | |:|     :::|
Db       1961 LPLCQRKEPKEASKESLAVTTKAIIT 1986


RESULT 15
DPOZ_MOUSE
ID   DPOZ_MOUSE     STANDARD;      PRT;  3122 AA.
```

```
AC   Q61493; Q9JMD6; Q9QWX6;
DT   30-MAY-2000 (Rel. 39, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   DNA polymerase zeta catalytic subunit (EC 2.7.7.7) (Seizure related
DE   protein 4).
GN   REV3L OR POLZ OR SEZ4.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129/Ola; TISSUE=Testis;
RX   MEDLINE=99202265; PubMed=10102037;
RA   Van Sloun P.P.H., Romeijn R.J., Eeken J.C.J.;
RT   "Molecular cloning, expression and chromosomal localisation of the
RT   mouse Rev3l gene, encoding the catalytic subunit of polymerase zeta.";
RL   Mutat. Res. 433:109-116(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Kajiwara K.;
RT   "Molecular analyses of Sez4 encoding murine homologue of yeast REV3 in
RT   brain neurons.";
RL   Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 2368-3122 FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Embryonic brain;
RX   MEDLINE=96216731; PubMed=8645260;
RA   Kajiwara K., Nagawawa H., Shimizu-Nishikawa K., Ookura T., Kimura M.,
RA   Sugaya E.;
RT   "Molecular characterization of seizure-related genes isolated by
RT   differential screening.";
RL   Biochem. Biophys. Res. Commun. 219:795-799(1996).
CC   -!- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC       + {DNA}(N).
CC   -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC   -!- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-B FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF083464; AAC98785.1; -.
DR   EMBL; AB031049; BAA90768.1; -.
DR   EMBL; D78644; BAA11461.1; -.
DR   PIR; T17202; T17202.
DR   MGD; MGI:1337131; Rev3l.
DR   InterPro; IPR006172; DNA_pol_B.
DR   InterPro; IPR006134; DNA_pol_B_dom.
DR   InterPro; IPR006133; DNA_pol_B_exo.
DR   InterPro; IPR004578; Pol2.
DR   Pfam; PF00136; DNA_pol_B; 1.
DR   Pfam; PF03104; DNA_pol_B_exo; 1.
DR   PRINTS; PR00106; DNAPOLB.
DR   SMART; SM00486; POLBc; 1.
DR   TIGRFAMs; TIGR00592; pol2; 1.
DR   PROSITE; PS00116; DNA_POLYMERASE_B; 1.
KW   Transferase; DNA-directed DNA polymerase; DNA replication;
KW   DNA-binding; DNA repair; Nuclear protein; Zinc-finger.
FT   ZN_FING    3034   3049       C4-TYPE (POTENTIAL).
FT   ZN_FING    3078   3096       C4-TYPE (POTENTIAL).
FT   CONFLICT     92     92       G -> A (IN REF. 2).
FT   CONFLICT    294    294       A -> T (IN REF. 2).
FT   CONFLICT    578    578       E -> Q (IN REF. 2).
FT   CONFLICT    609    609       R -> Q (IN REF. 2).
FT   CONFLICT   1278   1278       L -> P (IN REF. 2).
FT   CONFLICT   1298   1298       L -> F (IN REF. 2).
FT   CONFLICT   1416   1416       P -> L (IN REF. 2).
FT   CONFLICT   1848   1848       A -> T (IN REF. 2).
FT   CONFLICT   2368   2368       V -> G (IN REF. 3).
SQ   SEQUENCE   3122 AA;  350654 MW;  A39846CAF7365BA8 CRC64;

  Query Match             31.9%;  Score 45;  DB 1;  Length 3122;
  Best Local Similarity   33.3%;  Pred. No. 3.2e+02;
  Matches   10;  Conservative    8;  Mismatches   12;  Indels    0;  Gaps    0;

Qy          1 AVPIAQKSEPHSLSSEALMRRAVSLVTDST 30
              |:|  :| :||| :      ::|| :| |:
Db       1226 AIPADEKMKPHSEAELTPNHQSVSELTSSS 1255


Search completed: October  2, 2003, 09:37:22
Job time : 24 secs
```

THIS PAGE BLANK

GenCore version 5.1.6


OM protein - protein search, using sw model

Run on: October 2, 2003, 09:35:22 ; Search time 77 Seconds
(without alignments)
61.842 Million cell updates/sec

Title: US-09-939-293A-19_COPY_56_85
Perfect score: 141
Sequence: 1 AVPIAQKSEPHSLSSEALMRRAVSLVTDST 30

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 141 | 100.0 | 30 | 23 | AAU78435 | Inhibitor of apopt |
| 2 | 141 | 100.0 | 35 | 23 | AAU78439 | Inhibitor of apopt |
| 3 | 141 | 100.0 | 39 | 23 | AAU78436 | Inhibitor of apopt |
| 4 | 141 | 100.0 | 40 | 23 | AAU78430 | Inhibitor of apopt |
| 5 | 141 | 100.0 | 202 | 24 | ABG72302 | Human partial sequ |
| 6 | 141 | 100.0 | 227 | 21 | AAB54139 | Human pancreatic c |
| 7 | 141 | 100.0 | 239 | 21 | AAB26210 | Human caspase acti |
| 8 | 141 | 100.0 | 239 | 23 | AAU78447 | Inhibitor of apopt |
| 9 | 141 | 100.0 | 239 | 24 | ABP72164 | Human DIABLO/Smac. |
| 10 | 141 | 100.0 | 239 | 24 | ABB82743 | Human Smac polypep |
| 11 | 138 | 97.9 | 237 | 24 | ABG72301 | Mouse pro-apoptoti |
| 12 | 125 | 88.7 | 84 | 24 | ABG72303 | Rat partial sequen |
| 13 | 118 | 83.7 | 186 | 22 | AAB92922 | Human protein sequ |
| 14 | 96 | 68.1 | 20 | 23 | ABB76208 | Human smac (DIABLO |
| 15 | 70 | 49.6 | 15 | 24 | ABP71314 | Human Smac protein |
| 16 | 63 | 44.7 | 13 | 24 | ABG72314 | Human pro-apoptoti |
| 17 | 63 | 44.7 | 13 | 24 | ABG72316 | Human pro-apoptoti |
| 18 | 56 | 39.7 | 73 | 24 | ABG72304 | Flounder partial s |
| 19 | 50 | 35.5 | 396 | 22 | AAU51015 | Propionibacterium |
| 20 | 47.5 | 33.7 | 710 | 23 | ABP69647 | Human polypeptide |
| 21 | 47 | 33.3 | 944 | 23 | ABP35704 | Fungal ZBC protein |
| 22 | 47 | 33.3 | 2045 | 22 | ABB61941 | Drosophila melanog |
| 23 | 46 | 32.6 | 272 | 22 | AAG66438 | Human ATPase 30. |
| 24 | 46 | 32.6 | 284 | 23 | ABG79600 | Vernonia ribonucle |
| 25 | 46 | 32.6 | 312 | 22 | AAU27719 | Human full-length |
| 26 | 46 | 32.6 | 312 | 24 | ABP55411 | Human MDDT-20 prot |
| 27 | 46 | 32.6 | 336 | 22 | AAU27891 | Human contig polyp |
| 28 | 46 | 32.6 | 352 | 23 | ABP29265 | Streptococcus poly |
| 29 | 46 | 32.6 | 384 | 21 | AAG54380 | Arabidopsis thalia |
| 30 | 46 | 32.6 | 499 | 21 | AAG41345 | Arabidopsis thalia |
| 31 | 46 | 32.6 | 594 | 22 | ABB61195 | Drosophila melanog |
| 32 | 45.5 | 32.3 | 182 | 22 | ABU53172 | Human testes-deriv |
| 33 | 45 | 31.9 | 44 | 22 | ABB12208 | Human secreted pro |
| 34 | 45 | 31.9 | 124 | 22 | ABB57798 | Drosophila melanog |
| 35 | 45 | 31.9 | 173 | 22 | AAU48666 | Propionibacterium |
| 36 | 45 | 31.9 | 355 | 23 | ABP66105 | Bifidobacterium lo |
| 37 | 45 | 31.9 | 644 | 22 | ABG05466 | Novel human diagno |
| 38 | 45 | 31.9 | 809 | 23 | AAE15982 | Human cyclic nucle |
| 39 | 45 | 31.9 | 809 | 23 | AAE15983 | Human CNG3B protei |
| 40 | 45 | 31.9 | 809 | 23 | AAE15984 | Human CNG3B protei |
| 41 | 45 | 31.9 | 809 | 23 | AAE15985 | Human CNG3B protei |
| 42 | 45 | 31.9 | 809 | 23 | AAE15986 | Human CNG3B protei |
| 43 | 45 | 31.9 | 1018 | 22 | ABB62522 | Drosophila melanog |
| 44 | 45 | 31.9 | 2038 | 23 | AAE25098 | Human kinase and p |
| 45 | 45 | 31.9 | 2161 | 22 | AAM78959 | Human protein SEQ |

ALIGNMENTS

RESULT 1
AAU78435
ID AAU78435 standard; Peptide; 30 AA.
XX
AC AAU78435;
XX
DT 18-JUN-2002 (first entry)
XX
DE Inhibitor of apoptosis (IAP) protein Smac, mutant Smac-N30.
XX
KW Human; inhibitor of apoptosis; IAP; Smac; apoptosis; BID; BIR1; BIR2;
KW Bcl2 interacting domain; caspase; BIR domain; BIR3; gene therapy;
KW neoplastic cell; mutant; tumour.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200216418-A2.
XX
PD 28-FEB-2002.
XX
PF 24-AUG-2001; 2001WO-US26492.
XX
PR 24-AUG-2000; 2000US-227735P.
XX
PA (UYJE-) UNIV JEFFERSON THOMAS.
XX
PI Alnemri ES;
XX
DR WPI; 2002-304115/34.
XX
PT Novel Smac peptides and polynucleotides encoding the peptides, useful

PT for stimulating apoptosis in neoplastic or tumour cell which
PT overexpresses inhibitor of caspase, and for identifying apoptosis
PT modulating compounds -
XX
PS Example 3; Fig 7; 78pp; English.
XX
CC The invention relates to an isolated Smac peptide or polypeptide (I)
CC and an isolated nucleic acid (II) encoding (I). Also described is a
CC method of identifying a compound that inhibits apoptosis, comprising:
CC (a) separately contacting several cell populations expressing a
CC cytosolic Smac (a Smac isoform that begins with MKSDFYF sequence,
CC replacing the mitochondrial targeting sequence (residues 1-55 of (I)),
CC and residues 56-60 of (I)) and an inhibitor of BID (Bcl2 interacting
CC domain) with a compound to be tested for apoptotic inhibiting activity;
CC (b) incubating the cell populations with a direct stimulus of the cell
CC death pathway; and (c) measuring the specific apoptotic activity of the
CC cell populations, where inhibition of the specific apoptotic activity is
CC indicative that the compound is an inhibitor of apoptosis. (I) and (II)
CC are useful for inducing apoptosis in a cell. The Smac polypeptide and
CC polynucleotide are useful for stimulating apoptosis in a neoplastic or
CC tumour cell which overexpresses an inhibitor of caspase, where the
CC inhibitor inhibits activation or activity of caspase-3, caspase-7 or
CC caspase-9. Preferably, the cell overexpresses at least a portion of IAP.
CC (I) is useful for identifying an inhibitor or enhancer of a caspase-
CC mediated apoptosis which involves contacting a cell transformed or
CC transfected with a vector expressing (I) with a candidate inhibitor or
CC candidate enhancer; and detecting cell viability, where an increase in
CC cell viability indicates the presence of an inhibitor and a decrease in
CC cell viability indicates the presence of an enhancer. Optionally, the
CC method involves detecting the presence of large and small caspase
CC subunits after contacting cell transformed with the vector expressing
CC (I), with the candidate compound. A decrease in processing indicates the
CC presence of an inhibitor and an increase in the processing indicates the
CC presence of an enhancer. Preferably, the large and small subunits of
CC caspase-3, caspase-7 or caspase-9 are detected. (I) is also useful for
CC identifying a compound that inhibits Smac binding to Smac-binding
CC molecule (a portion of IAP e.g. a BIR domain such as BIR1, BIR2 or BIR3,
CC or a full-length IAP). (II) is useful in gene therapy techniques. The
CC present sequence represents the amino acid sequence of Smac mutant
CC Smac-N30.
XX
SQ Sequence 30 AA;

Query Match 100.0%; Score 141; DB 23; Length 30;
Best Local Similarity 100.0%; Pred. No. 2.2e-16;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AVPIAQKSEPHSLSSEALMRRAVSLVTDST 30
||||||||||||||||||||||||||||||
Db 1 AVPIAQKSEPHSLSSEALMRRAVSLVTDST 30

RESULT 2
AAU78439
ID AAU78439 standard; Peptide; 35 AA.
XX
AC AAU78439;
XX
DT 18-JUN-2002 (first entry)
XX
DE Inhibitor of apoptosis (IAP) protein Smac, peptide Smac-N35.
XX
KW Human; inhibitor of apoptosis; IAP; Smac; apoptosis; BID; BIR1; BIR2;
KW Bcl2 interacting domain; caspase; BIR domain; BIR3; gene therapy;
KW neoplastic cell; tumour.
XX
OS Homo sapiens.
XX
PN WO200216418-A2.
XX
PD 28-FEB-2002.
XX
PF 24-AUG-2001; 2001WO-US26492.
XX
PR 24-AUG-2000; 2000US-227735P.
XX
PA (UYJE-) UNIV JEFFERSON THOMAS.
XX
PI Alnemri ES;
XX
DR WPI; 2002-304115/34.
XX
PT Novel Smac peptides and polynucleotides encoding the peptides, useful
PT for stimulating apoptosis in neoplastic or tumour cell which
PT overexpresses inhibitor of caspase, and for identifying apoptosis
PT modulating compounds -
XX
PS Example 4; Page 47; 78pp; English.
XX
CC The invention relates to an isolated Smac peptide or polypeptide (I)
CC and an isolated nucleic acid (II) encoding (I). Also described is a
CC method of identifying a compound that inhibits apoptosis, comprising:
CC (a) separately contacting several cell populations expressing a
CC cytosolic Smac (a Smac isoform that begins with MKSDFYF sequence,
CC replacing the mitochondrial targeting sequence (residues 1-55 of (I)),
CC and residues 56-60 of (I)) and an inhibitor of BID (Bcl2 interacting
CC domain) with a compound to be tested for apoptotic inhibiting activity;
CC (b) incubating the cell populations with a direct stimulus of the cell
CC death pathway; and (c) measuring the specific apoptotic activity of the
CC cell populations, where inhibition of the specific apoptotic activity is
CC indicative that the compound is an inhibitor of apoptosis. (I) and (II)
CC are useful for inducing apoptosis in a cell. The Smac polypeptide and
CC polynucleotide are useful for stimulating apoptosis in a neoplastic or
CC tumour cell which overexpresses an inhibitor of caspase, where the
CC inhibitor inhibits activation or activity of caspase-3, caspase-7 or
CC caspase-9. Preferably, the cell overexpresses at least a portion of IAP.
CC (I) is useful for identifying an inhibitor or enhancer of a caspase-
CC mediated apoptosis which involves contacting a cell transformed or
CC transfected with a vector expressing (I) with a candidate inhibitor or
CC candidate enhancer; and detecting cell viability, where an increase in
CC cell viability indicates the presence of an inhibitor and a decrease in
CC cell viability indicates the presence of an enhancer. Optionally, the
CC method involves detecting the presence of large and small caspase
CC subunits after contacting cell transformed with the vector expressing
CC (I), with the candidate compound. A decrease in processing indicates the
CC presence of an inhibitor and an increase in the processing indicates the
CC presence of an enhancer. Preferably, the large and small subunits of
CC caspase-3, caspase-7 or caspase-9 are detected. (I) is also useful for
CC identifying a compound that inhibits Smac binding to Smac-binding
CC molecule (a portion of IAP e.g. a BIR domain such as BIR1, BIR2 or BIR3,
CC or a full-length IAP). (II) is useful in gene therapy techniques. The
CC present sequence represents the amino acid sequence of Smac peptide
CC Smac-N35.
XX
SQ Sequence 35 AA;

Query Match 100.0%; Score 141; DB 23; Length 35;
Best Local Similarity 100.0%; Pred. No. 2.7e-16;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AVPIAQKSEPHSLSSEALMRRAVSLVTDST 30
||||||||||||||||||||||||||||||
Db 1 AVPIAQKSEPHSLSSEALMRRAVSLVTDST 30

RESULT 3
AAU78436
ID AAU78436 standard; Peptide; 39 AA.
XX
AC AAU78436;
XX
DT 18-JUN-2002 (first entry)
XX
DE Inhibitor of apoptosis (IAP) protein Smac, mutant Smac-N39.
XX
KW Human; inhibitor of apoptosis; IAP; Smac; apoptosis; BID; BIR1; BIR2;

KW Bcl2 interacting domain; caspase; BIR domain; BIR3; gene therapy;
KW neoplastic cell; mutant; tumour.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200216418-A2.
XX
PD 28-FEB-2002.
XX
PF 24-AUG-2001; 2001WO-US26492.
XX
PR 24-AUG-2000; 2000US-227735P.
XX
PA (UYJE-) UNIV JEFFERSON THOMAS.
XX
PI Alnemri ES;
XX
DR WPI; 2002-304115/34.
XX
PT Novel Smac peptides and polynucleotides encoding the peptides, useful
PT for stimulating apoptosis in neoplastic or tumour cell which
PT overexpresses inhibitor of caspase, and for identifying apoptosis
PT modulating compounds -
XX
PS Example 3; Fig 7; 78pp; English.
XX
CC The invention relates to an isolated Smac peptide or polypeptide (I)
CC and an isolated nucleic acid (II) encoding (I). Also described is a
CC method of identifying a compound that inhibits apoptosis, comprising:
CC (a) separately contacting several cell populations expressing a
CC cytosolic Smac (a Smac isoform that begins with MKSDFYF sequence,
CC replacing the mitochondrial targeting sequence (residues 1-55 of (I)),
CC and residues 56-60 of (I)) and an inhibitor of BID (Bcl2 interacting
CC domain) with a compound to be tested for apoptotic inhibiting activity;
CC (b) incubating the cell populations with a direct stimulus of the cell
CC death pathway; and (c) measuring the specific apoptotic activity of the
CC cell populations, where inhibition of the specific apoptotic activity is
CC indicative that the compound is an inhibitor of apoptosis. (I) and (II)
CC are useful for inducing apoptosis in a cell. The Smac polypeptide and
CC polynucleotide are useful for stimulating apoptosis in a neoplastic or
CC tumour cell which overexpresses an inhibitor of caspase, where the
CC inhibitor inhibits activation or activity of caspase-3, caspase-7 or
CC caspase-9. Preferably, the cell overexpresses at least a portion of IAP.
CC (I) is useful for identifying an inhibitor or enhancer of a caspase-
CC mediated apoptosis which involves contacting a cell transformed or
CC transfected with a vector expressing (I) with a candidate inhibitor or
CC candidate enhancer; and detecting cell viability, where an increase in
CC cell viability indicates the presence of an inhibitor and a decrease in
CC cell viability indicates the presence of an enhancer. Optionally, the
CC method involves detecting the presence of large and small caspase
CC subunits after contacting cell transformed with the vector expressing
CC (I), with the candidate compound. A decrease in processing indicates the
CC presence of an inhibitor and an increase in the processing indicates the
CC presence of an enhancer. Preferably, the large and small subunits of
CC caspase-3, caspase-7 or caspase-9 are detected. (I) is also useful for
CC identifying a compound that inhibits Smac binding to Smac-binding
CC molecule (a portion of IAP e.g. a BIR domain such as BIR1, BIR2 or BIR3,
CC or a full-length IAP). (II) is useful in gene therapy techniques. The
CC present sequence represents the amino acid sequence of Smac mutant
CC Smac-N39.
XX
SQ Sequence 39 AA;

Query Match 100.0%; Score 141; DB 23; Length 39;
Best Local Similarity 100.0%; Pred. No. 3.1e-16;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AVPIAQKSEPHSLSSEALMRRAVSLVTDST 30
||||||||||||||||||||||||||||||
Db 1 AVPIAQKSEPHSLSSEALMRRAVSLVTDST 30

RESULT 4
AAU78430
ID AAU78430 standard; Peptide; 40 AA.
XX
AC AAU78430;
XX
DT 18-JUN-2002 (first entry)
XX
DE Inhibitor of apoptosis (IAP) protein Smac, N-terminal peptide.
XX
KW Human; inhibitor of apoptosis; IAP; Smac; apoptosis; BID; BIR1; BIR2;
KW Bcl2 interacting domain; caspase; BIR domain; BIR3; gene therapy;
KW neoplastic cell; tumour.
XX
OS Homo sapiens.
XX
PN WO200216418-A2.
XX
PD 28-FEB-2002.
XX
PF 24-AUG-2001; 2001WO-US26492.
XX
PR 24-AUG-2000; 2000US-227735P.
XX
PA (UYJE-) UNIV JEFFERSON THOMAS.
XX
PI Alnemri ES;
XX
DR WPI; 2002-304115/34.
XX
PT Novel Smac peptides and polynucleotides encoding the peptides, useful
PT for stimulating apoptosis in neoplastic or tumour cell which
PT overexpresses inhibitor of caspase, and for identifying apoptosis
PT modulating compounds -
XX
PS Example 3; Fig 7; 78pp; English.
XX
CC The invention relates to an isolated Smac peptide or polypeptide (I)
CC and an isolated nucleic acid (II) encoding (I). Also described is a
CC method of identifying a compound that inhibits apoptosis, comprising:
CC (a) separately contacting several cell populations expressing a
CC cytosolic Smac (a Smac isoform that begins with MKSDFYF sequence,
CC replacing the mitochondrial targeting sequence (residues 1-55 of (I)),
CC and residues 56-60 of (I)) and an inhibitor of BID (Bcl2 interacting
CC domain) with a compound to be tested for apoptotic inhibiting activity;
CC (b) incubating the cell populations with a direct stimulus of the cell
CC death pathway; and (c) measuring the specific apoptotic activity of the
CC cell populations, where inhibition of the specific apoptotic activity is
CC indicative that the compound is an inhibitor of apoptosis. (I) and (II)
CC are useful for inducing apoptosis in a cell. The Smac polypeptide and
CC polynucleotide are useful for stimulating apoptosis in a neoplastic or
CC tumour cell which overexpresses an inhibitor of caspase, where the
CC inhibitor inhibits activation or activity of caspase-3, caspase-7 or
CC caspase-9. Preferably, the cell overexpresses at least a portion of IAP.
CC (I) is useful for identifying an inhibitor or enhancer of a caspase-
CC mediated apoptosis which involves contacting a cell transformed or
CC transfected with a vector expressing (I) with a candidate inhibitor or
CC candidate enhancer; and detecting cell viability, where an increase in
CC cell viability indicates the presence of an inhibitor and a decrease in
CC cell viability indicates the presence of an enhancer. Optionally, the
CC method involves detecting the presence of large and small caspase
CC subunits after contacting cell transformed with the vector expressing
CC (I), with the candidate compound. A decrease in processing indicates the
CC presence of an inhibitor and an increase in the processing indicates the
CC presence of an enhancer. Preferably, the large and small subunits of
CC caspase-3, caspase-7 or caspase-9 are detected. (I) is also useful for
CC identifying a compound that inhibits Smac binding to Smac-binding
CC molecule (a portion of IAP e.g. a BIR domain such as BIR1, BIR2 or BIR3,
CC or a full-length IAP). (II) is useful in gene therapy techniques. The
CC present sequence represents the N-terminal amino acid sequence of Smac
CC protein.
XX
SQ Sequence 40 AA;

```
  Query Match          100.0%;  Score 141;  DB 23;  Length 40;
  Best Local Similarity  100.0%;  Pred. No. 3.2e-16;
  Matches   30;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AVPIAQKSEPHSLSSEALMRRAVSLVTDST 30
              ||||||||||||||||||||||||||||||
Db          1 AVPIAQKSEPHSLSSEALMRRAVSLVTDST 30


RESULT 5
ABG72302
ID   ABG72302 standard; Protein; 202 AA.
XX
AC   ABG72302;
XX
DT   29-JAN-2003  (first entry)
XX
DE   Human partial sequence for pro-apoptotic protein DIABLO.
XX
KW   Human; pro-apoptotic protein; DIABLO; cell death; apoptosis;
KW   inhibitor of apoptosis; IAP; cancer; vascular disease; hepatic disease;
KW   autoimmune disease; neurodegenerative disease; tissue damage;
KW   muscular tissue damage; heart attack; hepatic tissue damage;
KW   liver disease; immunogen.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   sig_peptide     1..25
FT                   /partial
FT   mat_peptide     26..202
FT                   /label= Mature_DIABLO
XX
PN   US2002110851-A1.
XX
PD   15-AUG-2002.
XX
PF   02-MAR-2001; 2001US-0798116.
XX
PR   02-MAR-2000; 2000AU-0005995.
XX
PA   (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
XX
PI   Verhagen AM,  Ekert PG,  Vaux DL;
XX
DR   WPI; 2003-074681/07.
XX
PT   New pro-apoptotic polypeptide, useful for screening for agents which
PT   modulate cell death and for treating conditions associated with cell
PT   death or apoptosis e.g. cancer  -
XX
PS   Disclosure; Fig 2E; 50pp; English.
XX
CC   The invention relates to an isolated pro-apoptotic polypeptide,
CC   designated DIABLO, or its biologically active fragment of 8 amino acids
CC   in length. Also included are the polynucleotide encoding DIABLO,
CC   expression vectors, transformed host cells, producing a biologically
CC   active fragment of DIABLO (by contacting an inhibitor of apoptosis (IAP)
CC   with a fragment of the polypeptide, and detecting a reduction in activity
CC   of the IAP), producing a natural or synthetic variant of DIABLO
CC   with cell death activity or which reduces IAP activity, an antigen-
CC   binding molecule that specifically binds to DIABLO or its fragment,
CC   detecting DIABLO in a biological sample (by contacting the sample
CC   with an IAP and detecting the presence of an IAP/DIABLO complex),
CC   modulating the death of a cell (by contacting a cell with an
CC   agent, which modulates the level and/or functional activity of a
CC   polypeptide), a composition for treatment/prophylaxis of a DIABLO related
CC   condition comprising an agent which reduces the level/activity of a
CC   polypeptide or DIABLO. DIABLO, or a nucleic acid encoding DIABLO, is
CC   useful for screening for an agent which modulates cell death. An
CC   antigen-binding molecule is useful for detecting DIABLO in a biological
CC   sample. The agent which modulates the level and/or functional activity of
CC   a polypeptide comprising mature or pro-human DIABLO polypeptide, is
CC   useful for the treatment and/or prophylaxis of a condition associated
CC   with expression or activation of DIABLO, such as cancer, vascular
CC   disease, hepatic disease, autoimmune disease and neurodegenerative
CC   disease, tissue damage or muscular tissue damage associated with heart
CC   attack, or hepatic tissue damage associated with a liver disease.
CC   DIABLO is also useful for treatment and/or prophylaxis of conditions
CC   associated with cell death or apoptosis. The present sequence
CC   represents partial human DIABLO.
XX
SQ   Sequence   202 AA;

  Query Match          100.0%;  Score 141;  DB 24;  Length 202;
  Best Local Similarity  100.0%;  Pred. No. 2.7e-15;
  Matches   30;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AVPIAQKSEPHSLSSEALMRRAVSLVTDST 30
              ||||||||||||||||||||||||||||||
Db         19 AVPIAQKSEPHSLSSEALMRRAVSLVTDST 48


RESULT 6
AAB54139
ID   AAB54139 standard; Protein; 227 AA.
XX
AC   AAB54139;
XX
DT   09-MAR-2001  (first entry)
XX
DE   Human pancreatic cancer antigen protein sequence SEQ ID NO:591.
XX
KW   Human; pancreas; pancreatic cancer; pancreatic cancer antigen;
KW   detection; diagnosis; identification; cytostatic; neuroprotective;
KW   nootropic; immunomodulatory; relaxant; contraceptive; gynaecological;
KW   antiinflammatory; cardiant; gene therapy; chromosome mapping;
KW   linkage analysis; tissue identification; tissue typing; forensic;
KW   neural; immune system; muscular; reproductive; gastrointestinal;
KW   pulmonary; cardiovascular; renal; proliferative.
XX
OS   Homo sapiens.
XX
PN   WO200055320-A1.
XX
PD   21-SEP-2000.
XX
PF   08-MAR-2000; 2000WO-US05989.
XX
PR   12-MAR-1999;   99US-0124270.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Ruben SM;
XX
DR   WPI; 2000-579444/54.
DR   N-PSDB; AAC98904.
XX
PT   New nucleic acid that is a pancreatic cancer antigen for preventing,
PT   treating, or ameliorating a medical condition, particular pancreatic
PT   cancer, or for use in assays for diagnosing a pathological condition -
XX
PS   Claim 11; Page 1027-1028; 1379pp; English.
XX
CC   AAC98773 to AAC99231 encode the human pancreatic cancer associated
CC   proteins, called pancreatic cancer antigens, given in AAB54008 to
CC   AAB54466. The human pancreatic cancer antigens have cytostatic,
CC   neuroprotective, nootropic, immunomodulatory, relaxant, contraceptive,
CC   gynaecological, cardiant and antiinflammatory activities, and can be used
CC   in gene therapy. The polynucleotide and proteins can be used for
CC   preventing, treating, or ameliorating a medical condition or in assays
CC   for diagnosing a pathological condition or a susceptibility to one in a
CC   subject. Binding partners to the proteins and the activity of the
CC   proteins can be identified. The pancreatic cancer antigens can be used to
CC   detect, treat or prevent pancreatic disorders, especially cancer.
```

CC Agonists and antagonists to the antigens can be screened for. The
CC pancreatic cancer antigen polynucleotides can be used to design nucleic
CC acid hybridisation probes that can be used in chromosome mapping, linkage
CC analysis, tissue identification and/or typing and a variety of forensic
CC and diagnostic methods. The proteins can be used to generate antibodies
CC which are used to purify, detect and target the polypeptides, including
CC both in vivo and in vitro diagnostic and therapeutic methods. The
CC proteins can be used to treat or prevent neural, immune system, muscular,
CC reproductive, gastrointestinal, pulmonary, cardiovascular, renal or
CC proliferative disorders. AAC99232 to AAC99240 and AAB54467 represent
CC sequences used in the exemplification of the present invention.
XX
SQ Sequence 227 AA;

Query Match 100.0%; Score 141; DB 21; Length 227;
Best Local Similarity 100.0%; Pred. No. 3.1e-15;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AVPIAQKSEPHSLSSEALMRRAVSLVTDST 30
||||||||||||||||||||||||||||||
Db 44 AVPIAQKSEPHSLSSEALMRRAVSLVTDST 73

RESULT 7
AAB26210
ID AAB26210 standard; Protein; 239 AA.
XX
AC AAB26210;
XX
DT 23-FEB-2001 (first entry)
XX
DE Human caspase activator Smac.
XX
KW Human; caspase activator; Smac; apoptosis; cancer; autoimmune disease;
KW neurodegenerative disease; mitochondria.
XX
OS Homo sapiens.
XX
PN US6110691-A.
XX
PD 29-AUG-2000.
XX
PF 06-JAN-2000; 2000US-0479309.
XX
PR 06-JAN-2000; 2000US-0479309.
XX
PA (TEXA ) UNIV TEXAS SYSTEM.
XX
PI Wang X, Du C;
XX
DR WPI; 2000-586350/55.
DR N-PSDB; AAA94860.
XX
PT Novel caspase regulatory polypeptides useful for screening binding
PT agents specific for the polypeptides which are useful for diagnosis and
PT also for treating apoptosis associated diseases -
XX
PS Claim 1; column 23-24; 16pp; English.
XX
CC The present sequence is the human Smac protein. Its coding sequence
CC was isolated by purifying the protein and searching a HeLa cell cDNA
CC library for sequences which bound to probes based upon it. Smac is a
CC mitochondrial protein which is released into the cytosol during
CC apoptosis, and acts as a caspase-3 activator. The protein and its coding
CC sequence can be used to modulate the expression and function of caspases
CC and their activators, and also can be used as drug targets and regulators
CC to promote or inhibit apoptosis in the treatment of cancer and autoimmune
CC and neurodegenerative diseases.
XX
SQ Sequence 239 AA;

Query Match 100.0%; Score 141; DB 21; Length 239;
Best Local Similarity 100.0%; Pred. No. 3.3e-15;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AVPIAQKSEPHSLSSEALMRRAVSLVTDST 30
||||||||||||||||||||||||||||||
Db 56 AVPIAQKSEPHSLSSEALMRRAVSLVTDST 85

RESULT 8
AAU78447
ID AAU78447 standard; Protein; 239 AA.
XX
AC AAU78447;
XX
DT 18-JUN-2002 (first entry)
XX
DE Inhibitor of apoptosis (IAP) protein Smac.
XX
KW Human; inhibitor of apoptosis; IAP; Smac; apoptosis; BID; BIR1; BIR2;
KW Bcl2 interacting domain; caspase; BIR domain; BIR3; gene therapy;
KW neoplastic cell; tumour.
XX
OS Homo sapiens.
XX
PN WO200216418-A2.
XX
PD 28-FEB-2002.
XX
PF 24-AUG-2001; 2001WO-US26492.
XX
PR 24-AUG-2000; 2000US-227735P.
XX
PA (UYJE-) UNIV JEFFERSON THOMAS.
XX
PI Alnemri ES;
XX
DR WPI; 2002-304115/34.
DR N-PSDB; ABK15451.
XX
PT Novel Smac peptides and polynucleotides encoding the peptides, useful
PT for stimulating apoptosis in neoplastic or tumour cell which
PT overexpresses inhibitor of caspase, and for identifying apoptosis
PT modulating compounds -
XX
PS Claim 36; Page 73-74; 78pp; English.
XX
CC The invention relates to an isolated Smac peptide or polypeptide (I)
CC and an isolated nucleic acid (II) encoding (I). Also described is a
CC method of identifying a compound that inhibits apoptosis, comprising:
CC (a) separately contacting several cell populations expressing a
CC cytosolic Smac (a Smac isoform that begins with MKSDFYF sequence,
CC replacing the mitochondrial targeting sequence (residues 1-55 of (I)),
CC and residues 56-60 of (I)) and an inhibitor of BID (Bcl2 interacting
CC domain) with a compound to be tested for apoptotic inhibiting activity;
CC (b) incubating the cell populations with a direct stimulus of the cell
CC death pathway; and (c) measuring the specific apoptotic activity of the
CC cell populations, where inhibition of the specific apoptotic activity is
CC indicative that the compound is an inhibitor of apoptosis. (I) and (II)
CC are useful for inducing apoptosis in a cell. The Smac polypeptide and
CC polynucleotide are useful for stimulating apoptosis in a neoplastic or
CC tumour cell which overexpresses an inhibitor of caspase, where the
CC inhibitor inhibits activation or activity of caspase-3, caspase-7 or
CC caspase-9. Preferably, the cell overexpresses at least a portion of IAP.
CC (I) is useful for identifying an inhibitor or enhancer of a caspase-
CC mediated apoptosis which involves contacting a cell transformed or
CC transfected with a vector expressing (I) with a candidate inhibitor or
CC candidate enhancer; and detecting cell viability, where an increase in
CC cell viability indicates the presence of an inhibitor and a decrease in
CC cell viability indicates the presence of an enhancer. Optionally, the
CC method involves detecting the presence of large and small caspase
CC subunits after contacting cell transformed with the vector expressing
CC (I), with the candidate compound. A decrease in processing indicates the
CC presence of an inhibitor and an increase in the processing indicates the
CC presence of an enhancer. Preferably, the large and small subunits of

```
CC   caspase-3, caspase-7 or caspase-9 are detected. (I) is also useful for
CC   identifying a compound that inhibits Smac binding to Smac-binding
CC   molecule (a portion of IAP e.g. a BIR domain such as BIR1, BIR2 or BIR3,
CC   or a full-length IAP). (II) is useful in gene therapy techniques. The
CC   present sequence represents the amino acid sequence of Smac protein.
XX
SQ   Sequence   239 AA;

  Query Match             100.0%;  Score 141;  DB 23;  Length 239;
  Best Local Similarity   100.0%;  Pred. No. 3.3e-15;
  Matches   30;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AVPIAQKSEPHSLSSEALMRRAVSLVTDST 30
              ||||||||||||||||||||||||||||||
Db         56 AVPIAQKSEPHSLSSEALMRRAVSLVTDST 85


RESULT 9
ABP72164
ID   ABP72164 standard; Protein; 239 AA.
XX
AC   ABP72164;
XX
DT   22-APR-2003  (first entry)
XX
DE   Human DIABLO/Smac.
XX
KW   Human; DIABLO/Smac; cell death; apoptosis;
KW   neurodegenerative disease; heart disease; cardiomyopathy; cardiant;
KW   neuroprotective; gene therapy.
XX
OS   Homo sapiens.
XX
PN   WO2003004606-A2.
XX
PD   16-JAN-2003.
XX
PF   03-JUL-2002; 2002WO-US21002.
XX
PR   03-JUL-2001; 2001US-0898158.
XX
PA   (UYCO ) UNIV COLUMBIA NEW YORK.
XX
PI   Troy CM,  Shelanski ML;
XX
DR   WPI; 2003-210351/20.
DR   N-PSDB; ABZ58109.
XX
PT   New nucleic acid encoding an inhibitor-of-apoptosis protein, useful for
PT   treating cancer, neurodegenerative disorder or cardiomyopathy  -
XX
PS   Disclosure; Fig 23A; 124pp; English.
XX
CC   The present sequence is the protein sequence for human DIABLO/Smac,
CC   an inhibitor of inhibitor-of-apoptosis (IAP) proteins.  The
CC   invention provides a nucleic acid, such as an antisense
CC   oligonucleotide, which specifically hybridises to a nucleic acid
CC   encoding a protein that induces cell death, especially APAF1, RAIDD
CC   or Diablo/SMAC.  A claimed method for inhibiting a cell's death
CC   (especially a neuronal cell's death) comprises contacting the cell
CC   with the nucleic acid under conditions permitting the nucleic acid
CC   to enter the cell, especially the use of a vector, liposome, or a
CC   mechanical or electrical means.  The method is used to a treat a
CC   neurodegenerative disorder, especially a brain disorder or central
CC   nervous system disorder, or a heart disorder, especially
CC   cardiomyopathy, in a human (all claimed).
XX
SQ   Sequence   239 AA;

  Query Match             100.0%;  Score 141;  DB 24;  Length 239;
  Best Local Similarity   100.0%;  Pred. No. 3.3e-15;
  Matches   30;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AVPIAQKSEPHSLSSEALMRRAVSLVTDST 30
              ||||||||||||||||||||||||||||||
Db         56 AVPIAQKSEPHSLSSEALMRRAVSLVTDST 85


RESULT 10
ABB82743
ID   ABB82743 standard; Protein; 239 AA.
XX
AC   ABB82743;
XX
DT   07-MAR-2003  (first entry)
XX
DE   Human Smac polypeptide.
XX
KW   Caspase-9; TUCAN; cancer; biomarker; cIAP2; Apaf1; Bcl-2; Smac;
KW   human.
XX
OS   Homo sapiens.
XX
PN   WO200290931-A2.
XX
PD   14-NOV-2002.
XX
PF   07-MAY-2002; 2002WO-US14487.
XX
PR   07-MAY-2001; 2001US-289223P.
PR   12-FEB-2002; 2002US-356934P.
XX
PA   (BURN-) BURNHAM INST.
XX
PI   Reed JC;
XX
DR   WPI; 2003-111999/10.
DR   N-PSDB; ABV75367.
XX
PT   Determining a prognosis for survival for a cancer patient, useful for
PT   determining if the patient is at risk for relapse, comprises measuring
PT   a level of TUCAN in a sample from the patient, and comparing it to a
PT   reference level  -
XX
PS   Examples; Page 151-153; 153pp; English.
XX
CC   The invention relates to determining a prognosis for survival for a
CC   cancer patient. The method involves (a) measuring a level of a tumour up-
CC   regulated CARD-containing antagonist of caspase-9 (TUCAN) in a neoplastic
CC   cell-containing sample from the cancer patient; and (b) comparing the
CC   level of TUCAN in the sample to a reference level of TUCAN, where a low
CC   level of TUCAN in the sample correlates with increased survival of the
CC   patient. Alternatively, the method involves measuring levels of TUCAN and
CC   one or more biomarkers selected from the group of cIAP2, Apaf1, Bcl-2, or
CC   Smac in a neoplastic cell-containing sample from the cancer patient. The
CC   method is useful for determining if the patient is at risk for relapse,
CC   or for determining a proper course of treatment for a patient with
CC   cancer. The method is also useful for monitoring the effectiveness of a
CC   course of treatment for a patient with cancer, e.g. colon cancer,
CC   gastrointestinal cancer, breast cancer, ovarian cancer, lung cancer,
CC   leukemia, CNS cancer, melanoma, prostate cancer, or renal cancer. The
CC   present sequence represents a human Smac polypeptide.
XX
SQ   Sequence   239 AA;

  Query Match             100.0%;  Score 141;  DB 24;  Length 239;
  Best Local Similarity   100.0%;  Pred. No. 3.3e-15;
  Matches   30;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AVPIAQKSEPHSLSSEALMRRAVSLVTDST 30
              ||||||||||||||||||||||||||||||
Db         56 AVPIAQKSEPHSLSSEALMRRAVSLVTDST 85


RESULT 11
ABG72301
```

```
ID   ABG72301 standard; Protein; 237 AA.
XX
AC   ABG72301;
XX
DT   29-JAN-2003  (first entry)
XX
DE   Mouse pro-apoptotic protein DIABLO.
XX
KW   Mouse; pro-apoptotic protein; DIABLO; cell death; apoptosis;
KW   inhibitor of apoptosis; IAP; cancer; vascular disease; hepatic disease;
KW   autoimmune disease; neurodegenerative disease; tissue damage;
KW   muscular tissue damage; heart attack; hepatic tissue damage;
KW   liver disease; immunogen.
XX
OS   Mus musculus.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..60
FT                   /label= Signal_peptide
FT   Peptide         1..8
FT                   /label= Immunogenic_fragment
FT                   /note= "This fragment is claimed in claim 4"
FT   Peptide         9..16
FT                   /label= Immunogenic_fragment
FT                   /note= "This fragment is claimed in claim 4"
FT   Peptide         17..24
FT                   /label= Immunogenic_fragment
FT                   /note= "This fragment is claimed in claim 4"
FT   Peptide         25..32
FT                   /label= Immunogenic_fragment
FT                   /note= "This fragment is claimed in claim 4"
FT   Peptide         33..40
FT                   /label= Immunogenic_fragment
FT                   /note= "This fragment is claimed in claim 4"
FT   Peptide         41..48
FT                   /label= Immunogenic_fragment
FT                   /note= "This fragment is claimed in claim 4"
FT   Peptide         49..56
FT                   /label= Immunogenic_fragment
FT                   /note= "This fragment is claimed in claim 4"
FT   Peptide         57..64
FT                   /label= Immunogenic_fragment
FT                   /note= "This fragment is claimed in claim 4"
FT   Peptide         65..72
FT                   /label= Immunogenic_fragment
FT                   /note= "This fragment is claimed in claim 4"
FT   Peptide         73..80
FT                   /label= Immunogenic_fragment
FT                   /note= "This fragment is claimed in claim 4"
FT   Peptide         81..88
FT                   /label= Immunogenic_fragment
FT                   /note= "This fragment is claimed in claim 4"
FT   Peptide         89..96
FT                   /label= Immunogenic_fragment
FT                   /note= "This fragment is claimed in claim 4"
FT   Peptide         97..104
FT                   /label= Immunogenic_fragment
FT                   /note= "This fragment is claimed in claim 4"
FT   Peptide         105..112
FT                   /label= Immunogenic_fragment
FT                   /note= "This fragment is claimed in claim 4"
FT   Peptide         113..120
FT                   /label= Immunogenic_fragment
FT                   /note= "This fragment is claimed in claim 4"
FT   Peptide         121..128
FT                   /label= Immunogenic_fragment
FT                   /note= "This fragment is claimed in claim 4"
FT   Peptide         129..136
FT                   /label= Immunogenic_fragment
FT                   /note= "This fragment is claimed in claim 4"
FT   Peptide         137..144
FT                   /label= Immunogenic_fragment
FT                   /note= "This fragment is claimed in claim 4"
FT   Peptide         145..152
FT                   /label= Immunogenic_fragment
FT                   /note= "This fragment is claimed in claim 4"
FT   Peptide         153..160
FT                   /label= Immunogenic_fragment
FT                   /note= "This fragment is claimed in claim 4"
FT   Protein         161..237
FT                   /label= Mature_DIABLO
FT                   /note= "This protein is claimed in claim 1"
FT   Peptide         161..168
FT                   /label= Immunogenic_fragment
FT                   /note= "This fragment is claimed in claim 4"
FT   Peptide         169..176
FT                   /label= Immunogenic_fragment
FT                   /note= "This fragment is claimed in claim 4"
FT   Peptide         177..184
FT                   /label= Immunogenic_fragment
FT                   /note= "This fragment is claimed in claim 4"
FT   Peptide         185..192
FT                   /label= Immunogenic_fragment
FT                   /note= "This fragment is claimed in claim 4"
FT   Peptide         193..200
FT                   /label= Immunogenic_fragment
FT                   /note= "This fragment is claimed in claim 4"
FT   Peptide         201..208
FT                   /label= Immunogenic_fragment
FT                   /note= "This fragment is claimed in claim 4"
FT   Peptide         209..216
FT                   /label= Immunogenic_fragment
FT                   /note= "This fragment is claimed in claim 4"
FT   Peptide         217..224
FT                   /label= Immunogenic_fragment
FT                   /note= "This fragment is claimed in claim 4"
FT   Peptide         225..232
FT                   /label= Immunogenic_fragment
FT                   /note= "This fragment is claimed in claim 4"
FT   Peptide         228..237
FT                   /label= Immunogenic_fragment
FT                   /note= "This fragment is claimed in claim 4"
XX
PN   US2002110851-A1.
XX
PD   15-AUG-2002.
XX
PF   02-MAR-2001; 2001US-0798116.
XX
PR   02-MAR-2000; 2000AU-0005995.
XX
PA   (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
XX
PI   Verhagen AM,  Ekert PG,  Vaux DL;
XX
DR   WPI; 2003-074681/07.
DR   N-PSDB; ABS57071.
XX
PT   New pro-apoptotic polypeptide, useful for screening for agents which
PT   modulate cell death and for treating conditions associated with cell
PT   death or apoptosis e.g. cancer  -
XX
PS   Claim 1; Fig 2E; 50pp; English.
XX
CC   The invention relates to an isolated pro-apoptotic polypeptide,
CC   designated DIABLO, or its biologically active fragment of 8 amino acids
CC   in length. Also included are the polynucleotide encoding DIABLO,
CC   expression vectors, transformed host cells, producing a biologically
CC   active fragment of DIABLO (by contacting an inhibitor of apoptosis (IAP)
CC   with a fragment of the polypeptide, and detecting a reduction in activity
CC   of the IAP), producing a natural or synthetic variant of DIABLO
CC   with cell death activity or which reduces IAP activity, an antigen-
CC   binding molecule that specifically binds to DIABLO or its fragment,
CC   detecting DIABLO in a biological sample (by contacting the sample
CC   with an IAP and detecting the presence of an IAP/DIABLO complex),
CC   modulating the death of a cell (by contacting a cell with an
```

```
CC   agent, which modulates the level and/or functional activity of a
CC   polypeptide), a composition for treatment/prophylaxis of a DIABLO related
CC   condition comprising an agent which reduces the level/activity of a
CC   polypeptide or DIABLO. DIABLO, or a nucleic acid encoding DIABLO, is
CC   useful for screening for an agent which modulates cell death. An
CC   antigen-binding molecule is useful for detecting DIABLO in a biological
CC   sample. The agent which modulates the level and/or functional activity of
CC   a polypeptide comprising mature or pro-human DIABLO polypeptide, is
CC   useful for the treatment and/or prophylaxis of a condition associated
CC   with expression or activation of DIABLO, such as cancer, vascular
CC   disease, hepatic disease, autoimmune disease and neurodegenerative
CC   disease, tissue damage or muscular tissue damage associated with heart
CC   attack, or hepatic tissue damage associated with a liver disease.
CC   DIABLO is also useful for treatment and/or prophylaxis of conditions
CC   associated with cell death or apoptosis. The present sequence
CC   represents mouse DIABLO.
XX
SQ   Sequence   237 AA;

  Query Match             97.9%;  Score 138;  DB 24;  Length 237;
  Best Local Similarity   96.7%;  Pred. No. 1.1e-14;
  Matches   29;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AVPIAQKSEPHSLSSEALMRRAVSLVTDST 30
              |||||||||||||:|||||||||||||||
Db         54 AVPIAQKSEPHSLSNEALMRRAVSLVTDST 83


RESULT 12
ABG72303
ID   ABG72303 standard; Protein; 84 AA.
XX
AC   ABG72303;
XX
DT   29-JAN-2003  (first entry)
XX
DE   Rat partial sequence for pro-apoptotic protein DIABLO.
XX
KW   Rat; pro-apoptotic protein; DIABLO; cell death; apoptosis;
KW   inhibitor of apoptosis; IAP; cancer; vascular disease; hepatic disease;
KW   autoimmune disease; neurodegenerative disease; tissue damage;
KW   muscular tissue damage; heart attack; hepatic tissue damage;
KW   liver disease; immunogen.
XX
OS   Rattus sp.
XX
PN   US2002110851-A1.
XX
PD   15-AUG-2002.
XX
PF   02-MAR-2001; 2001US-0798116.
XX
PR   02-MAR-2000; 2000AU-0005995.
XX
PA   (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
XX
PI   Verhagen AM,  Ekert PG,  Vaux DL;
XX
DR   WPI; 2003-074681/07.
XX
PT   New pro-apoptotic polypeptide, useful for screening for agents which
PT   modulate cell death and for treating conditions associated with cell
PT   death or apoptosis e.g. cancer  -
XX
PS   Disclosure; Page 35; 50pp; English.
XX
CC   The invention relates to an isolated pro-apoptotic polypeptide,
CC   designated DIABLO, or its biologically active fragment of 8 amino acids
CC   in length. Also included are the polynucleotide encoding DIABLO,
CC   expression vectors, transformed host cells, producing a biologically
CC   active fragment of DIABLO (by contacting an inhibitor of apoptosis (IAP)
CC   with a fragment of the polypeptide, and detecting a reduction in activity
CC   of the IAP), producing a natural or synthetic variant of DIABLO
CC   with cell death activity or which reduces IAP activity, an antigen-
CC   binding molecule that specifically binds to DIABLO or its fragment,
CC   detecting DIABLO in a biological sample (by contacting the sample
CC   with an IAP and detecting the presence of an IAP/DIABLO complex),
CC   modulating the death of a cell (by contacting a cell with an
CC   agent, which modulates the level and/or functional activity of a
CC   polypeptide), a composition for treatment/prophylaxis of a DIABLO related
CC   condition comprising an agent which reduces the level/activity of a
CC   polypeptide or DIABLO. DIABLO, or a nucleic acid encoding DIABLO, is
CC   useful for screening for an agent which modulates cell death. An
CC   antigen-binding molecule is useful for detecting DIABLO in a biological
CC   sample. The agent which modulates the level and/or functional activity of
CC   a polypeptide comprising mature or pro-human DIABLO polypeptide, is
CC   useful for the treatment and/or prophylaxis of a condition associated
CC   with expression or activation of DIABLO, such as cancer, vascular
CC   disease, hepatic disease, autoimmune disease and neurodegenerative
CC   disease, tissue damage or muscular tissue damage associated with heart
CC   attack, or hepatic tissue damage associated with a liver disease.
CC   DIABLO is also useful for treatment and/or prophylaxis of conditions
CC   associated with cell death or apoptosis. The present sequence
CC   represents partial rat DIABLO.
XX
SQ   Sequence   84 AA;

  Query Match             88.7%;  Score 125;  DB 24;  Length 84;
  Best Local Similarity   90.0%;  Pred. No. 4.2e-13;
  Matches   27;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 AVPIAQKSEPHSLSSEALMRRAVSLVTDST 30
              |||||||||| ||||:|||||||||||:||
Db         54 AVPIAQKSEPQSLSNEALMRRAVSLVTNST 83


RESULT 13
AAB92922
ID   AAB92922 standard; Protein; 186 AA.
XX
AC   AAB92922;
XX
DT   26-JUN-2001  (first entry)
XX
DE   Human protein sequence SEQ ID NO:11570.
XX
KW   Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX
OS   Homo sapiens.
XX
PN   EP1074617-A2.
XX
PD   07-FEB-2001.
XX
PF   28-JUL-2000; 2000EP-0116126.
XX
PR   29-JUL-1999;   99JP-0248036.
PR   27-AUG-1999;   99JP-0300253.
PR   11-JAN-2000;  2000JP-0118776.
PR   02-MAY-2000;  2000JP-0183767.
PR   09-JUN-2000;  2000JP-0241899.
XX
PA   (HELI-) HELIX RES INST.
XX
PI   Ota T,  Isogai T,  Nishikawa T,  Hayashi K,  Saito K,  Yamamoto J;
PI   Ishii S,  Sugiyama T,  Wakamatsu A,  Nagai K,  Otsuki T;
XX
DR   WPI; 2001-318749/34.
XX
PT   Primer sets for synthesizing polynucleotides, particularly the 5602
PT   full-length cDNAs defined in the specification, and for the detection
PT   and/or diagnosis of the abnormality of the proteins encoded by the
PT   full-length cDNAs -
XX
PS   Claim 8; SEQ ID 11570; 2537pp + CD ROM; English.
XX
```

CC The present invention describes primer sets for synthesising 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
SQ Sequence 186 AA;

Query Match 83.7%; Score 118; DB 22; Length 186;
Best Local Similarity 100.0%; Pred. No. 1.8e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 QKSEPHSLSSEALMRRAVSLVTDST 30
||||||||||||||||||||||||||
Db 8 QKSEPHSLSSEALMRRAVSLVTDST 32

RESULT 14
ABB76208
ID ABB76208 standard; Peptide; 20 AA.
XX
AC ABB76208;
XX
DT 09-AUG-2002 (first entry)
XX
DE Human smac (DIABLO) derived peptide.
XX
KW DIABLO; smac; inhibitor of apoptosis protein; IAP; apoptosis;
KW human; cancer; cytostatic.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Modified-site 20
FT /note= "optional C-terminal protecting group"
XX
PN WO200230959-A2.
XX
PD 18-APR-2002.
XX
PF 12-OCT-2001; 2001WO-US32121.
XX
PR 13-OCT-2000; 2000US-0687549.
XX
PA (ABBO ) ABBOTT LAB.
XX
PI Fesik SW, Meadows RP, Betz SP, Liu Z, Olejniczak ET, Sun C;
XX
DR WPI; 2002-444169/47.
XX
PT Novel peptide derived from wild-type human second mitochondria derived
PT activator of caspase protein useful for identifying candidate
PT substances to kill cancerous cells -
XX
PS Claim 5; Page 7; 26pp; English.
XX
CC The present sequence is a peptide derived from wild-type human
CC second mitochondria derived activator of caspase (smac), also known
CC as direct inhibitor of apoptosis binding protein with low pI
CC (DIABLO). The peptide is one of 12 claimed smac (DIABLO)-derived
CC peptides (see ABB76208-19) which bind to the Bir2 and Bir3 domain
CC of XIAP, an inhibitor of apoptosis protein (IAP) family member.
CC Kd values for Bir-3 and Bir-2 are 0.69 +/- 0.05 uM and 6.7 +/- 0.7
CC uM, respectively, for the present peptide, compared with 0.42 +/-
CC 0.02 uM and 2.3 +/- 0.3 uM, respectively, for full-length smac.
CC Modification of the N-terminal alanine destroys binding affinity to
CC XIAP. For example, N-terminal acetylation of the present peptide,
CC replacement of the N-terminal alanine with glycine, propionic acid
CC or isobityric acid all resulted in Kd values for Bir-3 and for Bir-2
CC of over 1,000 uM. The claimed peptides can be used to identify
CC candidate substances which induce or promote apoptosis in cells.
CC The assay involves determination of the ability of candidate
CC compounds to disrupt the binding interaction between a smac (DIABLO)
CC peptide and an IAP family member.
XX
SQ Sequence 20 AA;

Query Match 68.1%; Score 96; DB 23; Length 20;
Best Local Similarity 100.0%; Pred. No. 4.8e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AVPIAQKSEPHSLSSEALMR 20
||||||||||||||||||||
Db 1 AVPIAQKSEPHSLSSEALMR 20

RESULT 15
ABP71314
ID ABP71314 standard; peptide; 15 AA.
XX
AC ABP71314;
XX
DT 28-APR-2003 (first entry)
XX
DE Human Smac protein N-terminal fragment.
XX
KW Omi; HtrA2; serine protease; inhibitor of apoptosis protein; IAP;
KW caspase; apoptosis; cytostatic; immunosuppressive; neuroprotective;
KW vasotropic; gene therapy; Smac.
XX
OS Homo sapiens.
XX
PN WO2003006680-A2.
XX
PD 23-JAN-2003.
XX
PF 15-JUL-2002; 2002WO-US22658.
XX
PR 13-JUL-2001; 2001US-305378P.
PR 14-DEC-2001; 2001US-340163P.
XX
PA (UYJE-) UNIV JEFFERSON THOMAS.
XX
PI Alnemri ES;
XX
DR WPI; 2003-221760/21.
XX
PT New Omi nucleic acids and peptides that bind to an inhibitor of
PT apoptosis proteins, useful for regulating or altering caspase-mediated
PT apoptosis and for treating cancer, tumor, or autoimmune diseases -
XX
PS Example 2; Fig 6; 83pp; English.
XX
CC The invention relates to polynucleotides encoding an Omi (serine
CC protease) peptide or polypeptide. The Omi peptide specifically binds to a
CC portion of an Inhibitor of Apoptosis Protein (IAP). The Omi polypeptide
CC induces caspase-independent apoptosis, or fails to have serine protease
CC activity. The Omi peptides are useful for regulating or altering
CC apoptosis, specifically caspase-mediated apoptosis, and as immunogens for

CC raising antibodies. Enhancers of apoptosis are useful for treating
CC cancers, tumours or for destroying cells that mediate autoimmune
CC diseases. Compositions may also be used for the treatment of diseases
CC associated with inappropriate activation of apoptosis such as
CC neurodegenerative diseases and ischaemic injury. The antibodies can be
CC used in isolating Omi peptides, polypeptides and their variants, in
CC identifying molecules that interact with Omi peptides and polypeptides,
CC and in inhibiting or enhancing the biological activity of Omi peptides
CC and polypeptides. Sequences ABP71310-315 represent fragments of various
CC IAP-binding proteins, used to determine Omi as a IAP-binding protein.
XX
SQ Sequence 15 AA;

Query Match 49.6%; Score 70; DB 24; Length 15;
Best Local Similarity 93.3%; Pred. No. 7.7e-05;
Matches 14; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 AVPIAQKSEPHSLSS 15
              ||||||||||||||:
Db          1 AVPIAQKSEPHSLSN 15
```

Search completed: October 2, 2003, 09:36:50
Job time : 78 secs

GenCore version 5.1.6


OM protein - protein search, using sw model

Run on: October 2, 2003, 09:35:27 ; Search time 43 Seconds
(without alignments)
67.094 Million cell updates/sec

Title: US-09-939-293A-19_COPY_56_85
Perfect score: 141
Sequence: 1 AVPIAQKSEPHSLSSEALMRRAVSLVTDST 30

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 53 | 37.6 | 313 | 2 | T15529 | hypothetical prote |
| 2 | 53 | 37.6 | 321 | 2 | T15531 | hypothetical prote |
| 3 | 53 | 37.6 | 556 | 2 | S49754 | probable membrane |
| 4 | 52 | 36.9 | 630 | 2 | S77148 | hypothetical prote |
| 5 | 51 | 36.2 | 455 | 2 | S24458 | hypothetical prote |
| 6 | 50 | 35.5 | 256 | 2 | E75401 | hypothetical prote |
| 7 | 50 | 35.5 | 429 | 2 | JC4965 | elk1 protein - mou |
| 8 | 50 | 35.5 | 720 | 2 | T05616 | hypothetical prote |
| 9 | 49 | 34.8 | 338 | 2 | AD2754 | nitrogen regulatio |
| 10 | 49 | 34.8 | 338 | 2 | B97535 | nifr3-like protein |
| 11 | 48 | 34.0 | 210 | 2 | T15528 | hypothetical prote |
| 12 | 47 | 33.3 | 131 | 2 | F70663 | hypothetical prote |
| 13 | 47 | 33.3 | 390 | 2 | T01451 | protein kinase hom |
| 14 | 47 | 33.3 | 505 | 2 | H95946 | phosphate uptake A |
| 15 | 47 | 33.3 | 608 | 2 | B82635 | DNA polymerase III |
| 16 | 47 | 33.3 | 628 | 2 | S61160 | hypothetical prote |
| 17 | 47 | 33.3 | 944 | 1 | S48821 | probable membrane |
| 18 | 46 | 32.6 | 135 | 2 | H87410 | hypothetical prote |
| 19 | 46 | 32.6 | 205 | 2 | G82358 | conserved hypothet |
| 20 | 46 | 32.6 | 318 | 1 | D57987 | nrfD protein - Esc |
| 21 | 46 | 32.6 | 318 | 2 | G91260 | formate-dependent |
| 22 | 46 | 32.6 | 318 | 2 | C86101 | formate-dependent |
| 23 | 46 | 32.6 | 384 | 2 | F85439 | hypothetical prote |
| 24 | 46 | 32.6 | 525 | 2 | AF3601 | periplasmic oligop |
| 25 | 45.5 | 32.3 | 457 | 2 | T21344 | hypothetical prote |
| 26 | 45.5 | 32.3 | 531 | 2 | A84444 | hypothetical prote |
| 27 | 45 | 31.9 | 164 | 2 | E75293 | hypothetical prote |
| 28 | 45 | 31.9 | 315 | 2 | T40761 | conserved hypothet |
| 29 | 45 | 31.9 | 322 | 2 | T36841 | probable secreted |
| 30 | 45 | 31.9 | 356 | 2 | AE2784 | GGDEF family prote |
| 31 | 45 | 31.9 | 357 | 2 | F97563 | ggdef family prote |
| 32 | 45 | 31.9 | 474 | 2 | C75625 | hypothetical prote |
| 33 | 45 | 31.9 | 923 | 2 | A86334 | T20H2.17 protein - |
| 34 | 45 | 31.9 | 1211 | 2 | T08540 | hypothetical prote |
| 35 | 45 | 31.9 | 1888 | 2 | T14273 | zinc finger protei |
| 36 | 45 | 31.9 | 2515 | 2 | A41519 | posterior-group pr |
| 37 | 45 | 31.9 | 2643 | 2 | T29149 | hypothetical prote |
| 38 | 45 | 31.9 | 3122 | 2 | T17202 | DNA-directed DNA p |
| 39 | 44.5 | 31.6 | 4151 | 2 | T13734 | groovin gene prote |
| 40 | 44 | 31.2 | 170 | 2 | G69541 | conserved hypothet |
| 41 | 44 | 31.2 | 224 | 2 | T43331 | clathrin light cha |
| 42 | 44 | 31.2 | 229 | 2 | T40789 | clathrin light cha |
| 43 | 44 | 31.2 | 270 | 2 | F86177 | protein F19P19.1 [ |
| 44 | 44 | 31.2 | 363 | 2 | B84565 | probable spliceoso |
| 45 | 44 | 31.2 | 386 | 2 | A96625 | hypothetical prote |

ALIGNMENTS

RESULT 1
T15529
hypothetical protein C17C3.7 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C;Accession: T15529
R;Du, Z.
submitted to the EMBL Data Library, November 1995
A;Description: The sequence of C. elegans cosmid C17C3.
A;Reference number: Z18366
A;Accession: T15529
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-313 <DUZ>
A;Cross-references: EMBL:U41279; NID:g1086905; PID:g1086908; PIDN:AAB52691.1; GSPDB:G
A;Experimental source: strain Bristol N2; clone C17C3
C;Genetics:
A;Gene: CESP:C17C3.7
A;Map position: 2
A;Introns: 45/3; 98/2; 175/2

Query Match 37.6%; Score 53; DB 2; Length 313;
Best Local Similarity 50.0%; Pred. No. 2.3;
Matches 12; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

```
Qy          4 IAQKSEPHSLSSEALMRRAVSLVT 27
              | ||||   :| | :: | | |||
Db         88 IVQKSEEEHISQEVVLFRIVKLVT 111
```

RESULT 2
T15531
hypothetical protein C17C3.10 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C;Accession: T15531
R;Du, Z.
submitted to the EMBL Data Library, November 1995
A;Description: The sequence of C. elegans cosmid C17C3.
A;Reference number: Z18366
A;Accession: T15531
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-321 <DUZ>
A;Cross-references: EMBL:U41279; NID:g1086905; PID:g1086910; PIDN:AAB52693.1; GSPDB:G
A;Experimental source: strain Bristol N2; clone C17C3
C;Genetics:
A;Gene: CESP:C17C3.10
A;Map position: 2
A;Introns: 16/3; 38/3; 53/3; 106/2; 183/2

Query Match 37.6%; Score 53; DB 2; Length 321;

```
  Best Local Similarity   50.0%;  Pred. No. 2.4;
  Matches   12;  Conservative    3;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          4 IAQKSEPHSLSSEALMRRAVSLVT 27
              | ||||   :| | :: | | |||
Db         96 IVQKSEEEHISQEVVLFRIVKLVT 119


RESULT 3
S49754
probable membrane protein YML023c - yeast (Saccharomyces cerevisiae)
C;Species: Saccharomyces cerevisiae
C;Date: 13-Jan-1995 #sequence_revision 10-Feb-1995 #text_change 19-Apr-2002
C;Accession: S49754
R;Badcock, K.; Churcher, C.
submitted to the EMBL Data Library, November 1994
A;Reference number: S49741
A;Accession: S49754
A;Molecule type: DNA
A;Residues: 1-556 <BAD>
A;Cross-references: EMBL:Z46659; NID:g575680; PID:g575696; GSPDB:GN00013; MIPS:YML023c
C;Genetics:
A;Gene: MIPS:YML023c
A;Cross-references: SGD:S0004485
A;Map position: 13L
C;Superfamily: Saccharomyces cerevisiae probable membrane protein YML023c
C;Keywords: transmembrane protein
F;61-77/Domain: transmembrane #status predicted <TM1>
F;482-498/Domain: transmembrane #status predicted <TM2>

  Query Match             37.6%;  Score 53;  DB 2;  Length 556;
  Best Local Similarity   41.7%;  Pred. No. 4.5;
  Matches   10;  Conservative    7;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          3 PIAQKSEPHSLSSEALMRRAVSLV 26
              || : ::|:: | | | |||: |:
Db         85 PILRANDPYNTSRETLSRRALKLL 108


RESULT 4
S77148
hypothetical protein slr1888 - Synechocystis sp. (strain PCC 6803)
C;Species: Synechocystis sp.
A;Variety: PCC 6803
C;Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 08-Oct-1999
C;Accession: S77148
R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
o, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A;Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
s.
A;Reference number: S74322; MUID:97061201; PMID:8905231
A;Accession: S77148
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-630 <KAN>
A;Cross-references: EMBL:D90908; GB:AB001339; NID:g1652725; PIDN:BAA17706.1; PID:d101843
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

  Query Match             36.9%;  Score 52;  DB 2;  Length 630;
  Best Local Similarity   40.6%;  Pred. No. 7.4;
  Matches   13;  Conservative    6;  Mismatches    7;  Indels    6;  Gaps    1;

Qy          3 PIAQKSEP------HSLSSEALMRRAVSLVTD 28
              |:    ||      :||| |:|:|| :|: ||
Db        479 PVKFNDEPNLKNFFYSLSDESLVRRFMSVRTD 510


RESULT 5
S24458
hypothetical protein R08D7.2 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 19-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 09-Sep-1997
C;Accession: S41037; S24458
R;Ainscough, R.; Hawkins, T.
submitted to the EMBL Data Library, May 1992
A;Reference number: S41036
A;Accession: S41037
A;Molecule type: DNA
A;Residues: 1-455 <AIN>
A;Cross-references: EMBL:Z12017; NID:g6812; PID:g6814
R;Sulston, J.; Du, Z.; Thomas, K.; Wilson, R.; Hillier, L.; Staden, R.; Halloran, N.;
awkins, T.; Ainscough, R.; Waterston, R.
Nature 356, 37-41, 1992
A;Title: The C. elegans genome sequencing project: a beginning.
A;Reference number: S24457; MUID:92168156; PMID:1538779
A;Contents: annotation
C;Genetics:
A;Introns: 37/3; 128/1; 165/3; 291/2; 357/2; 407/2

  Query Match             36.2%;  Score 51;  DB 2;  Length 455;
  Best Local Similarity   48.1%;  Pred. No. 7.2;
  Matches   13;  Conservative    5;  Mismatches    7;  Indels    2;  Gaps    1;

Qy          3 PIAQKSEPHSLSSEALMRRAVSLVTDS 29
              |||:|||| : |:     :: |||||
Db        141 PIAEKSEP--VDSQFKKEPSIWLVTDS 165


RESULT 6
E75401
hypothetical protein - Deinococcus radiodurans (strain R1)
C;Species: Deinococcus radiodurans
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
C;Accession: E75401
R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J
, M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.;
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A;Reference number: A75250; MUID:20036896; PMID:10567266
A;Accession: E75401
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-256 <WHI>
A;Cross-references: GB:AE001984; GB:AE000513; NID:g6459135; PIDN:AAF10959.1; PID:g645
A;Experimental source: strain R1
C;Genetics:
A;Gene: DR1380
A;Map position: 1

  Query Match             35.5%;  Score 50;  DB 2;  Length 256;
  Best Local Similarity   38.1%;  Pred. No. 5.3;
  Matches    8;  Conservative    6;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          2 VPIAQKSEPHSLSSEALMRRA 22
              :|:| : :|| |:   |:| |
Db        188 IPLALRQDPHRLAQSGLVRAA 208


RESULT 7
JC4965
elk1 protein - mouse
C;Species: Mus musculus (house mouse)
C;Date: 31-Dec-1996 #sequence_revision 31-Dec-1996 #text_change 16-Jul-1999
C;Accession: JC4965; I48339; I48340; S54908
R;Grevin, D.; Ung, S.; Denhez, F.; Dehem, M.; Quatannens, B.; Begue, A.; Stehelin, D.
Gene 174, 185-188, 1996
A;Title: Structure and organization of the mouse elk1 gene.
A;Reference number: JC4965; MUID:97017146; PMID:8863747
A;Accession: JC4965
A;Molecule type: mRNA
A;Residues: 1-429 <GRE>
A;Cross-references: EMBL:X87257; NID:g836634; PIDN:CAA60715.1; PID:g836635
A;Experimental source: embyro
R;Giovane, A.; Pintzas, A.; Maira, S.M.; Sobieszczuk, P.; Wasylyk, B.
```

```
Genes Dev. 8, 1502-1513, 1994
A;Title: Net, a new ets transcription factor that is activated by Ras.
A;Reference number: A53837; MUID:95047310; PMID:7958835
A;Accession: I48339
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-429 <RES>
A;Cross-references: EMBL:X87257; NID:g836634; PIDN:CAA60715.1; PID:g836635
A;Accession: I48340
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 5-132,'T',134-224 <RE2>
A;Cross-references: EMBL:Z36939; NID:g535922; PIDN:CAA85391.1; PID:g535923
C;Comment: This protein belongs to the subfamily of ternary complex factor (TCF) which b
C;Genetics:
A;Gene: elk1
A;Introns: 70/3; 219/3; 363/3; 397/3
C;Superfamily: elk-1 transforming protein; ets DNA-binding domain homology
F;7-86/Domain: ets DNA-binding domain homology <ETS>

  Query Match             35.5%;  Score 50;  DB 2;  Length 429;
  Best Local Similarity   38.5%;  Pred. No. 9.7;
  Matches   10;  Conservative    8;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          5 AQKSEPHSLSSEALMRRAVSLVTDST 30
              | |:||   :|| |: |  :::|::|
Db        270 AVKAEPEVSASEGLLARLPAILTENT 295


RESULT 8
T05616
hypothetical protein F20D10.10 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 23-Sep-2002
C;Accession: T05616
R;Bevan, M.; Wedler, H.; Kutzner, M.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; Mayer, K.F
submitted to the Protein Sequence Database, February 1999
A;Reference number: Z15420
A;Accession: T05616
A;Molecule type: DNA
A;Residues: 1-720 <BEV>
A;Cross-references: EMBL:AL035538
A;Experimental source: cultivar Columbia; BAC clone F20D10
C;Genetics:
A;Map position: 4
A;Introns: 142/3
A;Note: F20D10.10
C;Superfamily: retroelement pol polyprotein; RING finger homology
F;146-201/Domain: RING finger homology <RRN>

  Query Match             35.5%;  Score 50;  DB 2;  Length 720;
  Best Local Similarity   33.3%;  Pred. No. 18;
  Matches   12;  Conservative    7;  Mismatches    7;  Indels   10;  Gaps    1;

Qy          3 PIAQKSEP----------HSLSSEALMRRAVSLVTD 28
              |: |||||           |:  |:||: :: |:|
Db        676 PVVQKSEPLTPTSAWRAAERLAKVAIMRKHMNRVSD 711


RESULT 9
AD2754
nitrogen regulation protein nifR [imported] - Agrobacterium tumefaciens (strain C58, Dup
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
C;Accession: AD2754
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L
erage, G.; Gillet, W.; Grant, C.; Guenthner, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; MUID:21608550; PMID:11743193
A;Accession: AD2754
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-338 <KUR>
A;Cross-references: GB:AE008688; PIDN:AAL42450.1; PID:g17739864; GSPDB:GN00186
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: nifR
A;Map position: circular chromosome
C;Superfamily: conserved hypothetical protein HI0979

  Query Match             34.8%;  Score 49;  DB 2;  Length 338;
  Best Local Similarity   41.9%;  Pred. No. 10;
  Matches   13;  Conservative    5;  Mismatches    9;  Indels    4;  Gaps    1;

Qy          2 VPIAQKSEPHSLSSEALMRRAV----SLVTD 28
              :|: : ||| |: |  :  |||    | |||
Db          6 LPLPELSEPFSIGSVTIRNRAVLAPMSGVTD 36


RESULT 10
B97535
nifr3-like protein [imported] - Agrobacterium tumefaciens (strain C58, Cereon)
C;Species: Agrobacterium tumefaciens
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 18-Nov-2002
C;Accession: B97535
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Gold
 A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz,
Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium
A;Reference number: A97359; MUID:21608551; PMID:11743194
A;Accession: B97535
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-338 <KUR>
A;Cross-references: GB:AE007869; PIDN:AAK87235.1; PID:g15156519; GSPDB:GN00169
C;Genetics:
A;Gene: AGR_C_2661
A;Map position: circular chromosome
C;Superfamily: conserved hypothetical protein HI0979

  Query Match             34.8%;  Score 49;  DB 2;  Length 338;
  Best Local Similarity   41.9%;  Pred. No. 10;
  Matches   13;  Conservative    5;  Mismatches    9;  Indels    4;  Gaps    1;

Qy          2 VPIAQKSEPHSLSSEALMRRAV----SLVTD 28
              :|: : ||| |: |  :  |||    | |||
Db          6 LPLPELSEPFSIGSVTIRNRAVLAPMSGVTD 36


RESULT 11
T15528
hypothetical protein C17C3.8 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C;Accession: T15528
R;Du, Z.
submitted to the EMBL Data Library, November 1995
A;Description: The sequence of C. elegans cosmid C17C3.
A;Reference number: Z18366
A;Accession: T15528
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-210 <DUZ>
A;Cross-references: EMBL:U41279; NID:g1086905; PID:g1086907; PIDN:AAB52692.1; GSPDB:G
A;Experimental source: strain Bristol N2; clone C17C3
C;Genetics:
A;Gene: CESP:C17C3.8
A;Map position: 2
A;Introns: 50/2; 119/2

  Query Match             34.0%;  Score 48;  DB 2;  Length 210;
  Best Local Similarity   45.8%;  Pred. No. 8.6;
```

```
  Matches   11;  Conservative    4;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          4 IAQKSEPHSLSSEALMRRAVSLVT 27
              | ||||   || | :: | | |::
Db         40 IVQKSESEKLSKEEVLFRIVKLLS 63


RESULT 12
F70663
hypothetical protein Rv1838c - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C;Accession: F70663
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Sqares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: F70663
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-131 <COL>
A;Cross-references: GB:Z83859; GB:AL123456; NID:g3261678; PIDN:CAB06116.1; PID:g1781209
A;Experimental source: strain H37Rv
C;Genetics:
A;Gene: Rv1838c
C;Superfamily: conserved hypothetical protein MJ0974

  Query Match             33.3%;  Score 47;  DB 2;  Length 131;
  Best Local Similarity   42.3%;  Pred. No. 7.1;
  Matches   11;  Conservative    5;  Mismatches    6;  Indels    4;  Gaps    1;

Qy          8 SEPHSLSSEALMRRAVS----LVTDS 29
              | || | :: |:  |:|    ||||:
Db         16 SHPHKLDAQRLLESALSGGERLVTDA 41


RESULT 13
T01451
protein kinase homolog F24O1.13 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 23-Mar-2001
C;Accession: T01451
R;Shinn, P.; Buehler, E.; Dewar, K.; Feng, J.; Kim, C.; Li, Y.; Sun, H.; Conway, A.; Con
eologis, A.; Ecker, J.R.
submitted to the EMBL Data Library, January 1998
A;Description: Genomic sequence for Arabidopsis thaliana BAC F24O1.
A;Reference number: Z14211
A;Accession: T01451
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-390 <SHI>
A;Cross-references: EMBL:AC003113; NID:g2689438; PID:g2781357; GSPDB:GN00059; ATSP:F24O1
C;Genetics:
A;Gene: ATSP:F24O1.13
A;Map position: 1
A;Introns: 149/3; 301/3
C;Superfamily: kinase-related transforming protein; protein kinase homology

  Query Match             33.3%;  Score 47;  DB 2;  Length 390;
  Best Local Similarity   47.4%;  Pred. No. 25;
  Matches    9;  Conservative    5;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          7 KSEPHSLSSEALMRRAVSL 25
              | ||:||| | ::| |: :
Db        178 KKEPYSLSIETVLRLALDI 196


RESULT 14
H95946
phosphate uptake ABC transporter permease protein phoT [imported] - Sinorhizobium melild
C;Species: Sinorhizobium meliloti
C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
C;Accession: H95946
R;Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Her
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A;Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing e
A;Reference number: A95842; MUID:21396508; PMID:11481431
A;Accession: H95946
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-505 <KUR>
A;Cross-references: GB:AL591985; PIDN:CAC49240.1; PID:g15140726; GSPDB:GN00167
A;Experimental source: strain 1021, megaplasmid pSymB
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubl
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.
 L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelau
hebault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh,
A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A;Reference number: A96039; MUID:21368234; PMID:11474104
A;Contents: annotation
C;Genetics:
A;Gene: phoT; SMb21174
A;Genome: plasmid

  Query Match             33.3%;  Score 47;  DB 2;  Length 505;
  Best Local Similarity   50.0%;  Pred. No. 34;
  Matches    9;  Conservative    3;  Mismatches    6;  Indels    0;  Gaps    0;

Qy         10 PHSLSSEALMRRAVSLVT 27
              || | | |  ||: :|:|
Db         20 PHLLQSSAAKRRSTALIT 37


RESULT 15
B82635
DNA polymerase III subunit XF1807 [imported] - Xylella fastidiosa (strain 9a5c)
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 15-Sep-2000
C;Accession: B82635
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Seq
Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: B82635
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-608 <SIM>
A;Cross-references: GB:AE004002; GB:AE003849; NID:g9106881; PIDN:AAF84614.1; GSPDB:GN
A;Experimental source: strain 9a5c
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer
as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Fr
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; La
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.
, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri,
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawa
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silv
M.; Tsuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF1807
C;Superfamily: DNA-directed DNA polymerase III gamma chain

  Query Match             33.3%;  Score 47;  DB 2;  Length 608;
  Best Local Similarity   38.2%;  Pred. No. 43;
  Matches   13;  Conservative    4;  Mismatches    7;  Indels   10;  Gaps    1;
```

```
Qy          6 QKSEPHSLSSEALM----------RRAVSLVTDS 29
              ||||   |: || |          ::|  :||||
Db        569 QKSERQQLAEEAFMSDHVVQHLIYQQAAKVVTDS 602
```

Search completed: October 2, 2003, 09:40:07
Job time : 45 secs

GenCore version 5.1.6


OM protein - protein search, using sw model

Run on: October 2, 2003, 09:35:22 ; Search time 103 Seconds
(without alignments)
75.161 Million cell updates/sec

Title: US-09-939-293A-19_COPY_56_85
Perfect score: 141
Sequence: 1 AVPIAQKSEPHSLSSEALMRRAVSLVTDST 30

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 130 | 92.2 | 157 | 11 | Q8R1D8 | Q8r1d8 mus musculu |
| 2 | 56 | 39.7 | 224 | 10 | Q9AX95 | Q9ax95 oryza sativ |
| 3 | 53 | 37.6 | 313 | 5 | Q18054 | Q18054 caenorhabdi |
| 4 | 53 | 37.6 | 321 | 5 | Q18056 | Q18056 caenorhabdi |
| 5 | 52 | 36.9 | 181 | 16 | Q8DF79 | Q8df79 vibrio vuln |
| 6 | 52 | 36.9 | 207 | 16 | Q8DD72 | Q8dd72 vibrio vuln |
| 7 | 52 | 36.9 | 630 | 16 | P73661 | P73661 synechocyst |
| 8 | 52 | 36.9 | 862 | 13 | Q8AVG0 | Q8avg0 xenopus lae |
| 9 | 51 | 36.2 | 196 | 2 | Q9ZIG9 | Q9zig9 pseudomonas |
| 10 | 50.5 | 35.8 | 928 | 16 | Q8DC68 | Q8dc68 vibrio vuln |
| 11 | 50 | 35.5 | 256 | 16 | Q9RUK4 | Q9ruk4 deinococcus |
| 12 | 50 | 35.5 | 720 | 10 | Q9SZJ1 | Q9szj1 arabidopsis |
| 13 | 50 | 35.5 | 739 | 10 | Q8GYZ4 | Q8gyz4 arabidopsis |
| 14 | 50 | 35.5 | 981 | 4 | Q9BYT9 | Q9byt9 homo sapien |
| 15 | 50 | 35.5 | 1003 | 16 | Q8EUZ4 | Q8euz4 mycoplasma |
| 16 | 49 | 34.8 | 338 | 16 | Q8UFF3 | Q8uff3 agrobacteri |
| 17 | 49 | 34.8 | 561 | 5 | Q8SV66 | Q8sv66 encephalito |
| 18 | 48.5 | 34.4 | 382 | 16 | Q98MH8 | Q98mh8 rhizobium l |
| 19 | 48 | 34.0 | 74 | 9 | Q9XJS5 | Q9xjs5 bacteriopha |
| 20 | 48 | 34.0 | 210 | 5 | Q18053 | Q18053 caenorhabdi |
| 21 | 48 | 34.0 | 268 | 16 | Q8XYJ1 | Q8xyj1 ralstonia s |
| 22 | 48 | 34.0 | 343 | 16 | Q8PI58 | Q8pi58 xanthomonas |
| 23 | 48 | 34.0 | 343 | 16 | Q8P6V8 | Q8p6v8 xanthomonas |
| 24 | 48 | 34.0 | 477 | 13 | Q9W696 | Q9w696 xenopus lae |
| 25 | 47.5 | 33.7 | 684 | 4 | Q9HCM6 | Q9hcm6 homo sapien |
| 26 | 47.5 | 33.7 | 700 | 3 | Q8X007 | Q8x007 neurospora |
| 27 | 47 | 33.3 | 282 | 16 | Q8XZX2 | Q8xzx2 ralstonia s |
| 28 | 47 | 33.3 | 333 | 17 | Q8TH73 | Q8th73 methanosarc |
| 29 | 47 | 33.3 | 415 | 10 | Q9MAV2 | Q9mav2 arabidopsis |
| 30 | 47 | 33.3 | 505 | 16 | Q52909 | Q52909 rhizobium m |
| 31 | 47 | 33.3 | 516 | 11 | Q8CA49 | Q8ca49 mus musculu |
| 32 | 47 | 33.3 | 608 | 16 | Q9PCH3 | Q9pch3 xylella fas |
| 33 | 47 | 33.3 | 628 | 3 | Q06344 | Q06344 saccharomyc |
| 34 | 47 | 33.3 | 2016 | 5 | Q9W444 | Q9w444 drosophila |
| 35 | 46.5 | 33.0 | 434 | 16 | Q8XRD2 | Q8xrd2 ralstonia s |
| 36 | 46 | 32.6 | 120 | 10 | Q9ST89 | Q9st89 oryza sativ |
| 37 | 46 | 32.6 | 135 | 16 | Q9A8Q0 | Q9a8q0 caulobacter |
| 38 | 46 | 32.6 | 205 | 16 | Q9KVK7 | Q9kvk7 vibrio chol |
| 39 | 46 | 32.6 | 258 | 4 | Q8IXN3 | Q8ixn3 homo sapien |
| 40 | 46 | 32.6 | 312 | 11 | Q8BUZ6 | Q8buz6 mus musculu |
| 41 | 46 | 32.6 | 318 | 4 | Q9P0R9 | Q9p0r9 homo sapien |
| 42 | 46 | 32.6 | 318 | 16 | Q8X4L4 | Q8x4l4 escherichia |
| 43 | 46 | 32.6 | 319 | 4 | Q9P0P6 | Q9p0p6 homo sapien |
| 44 | 46 | 32.6 | 352 | 2 | O85471 | O85471 streptococc |
| 45 | 46 | 32.6 | 352 | 16 | Q9A0K4 | Q9a0k4 streptococc |

ALIGNMENTS

RESULT 1
Q8R1D8
ID Q8R1D8 PRELIMINARY; PRT; 157 AA.
AC Q8R1D8;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Similar to RIKEN cDNA 0610041G12 gene.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Eye;
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; BC024780; AAH24780.1; -.
SQ SEQUENCE 157 AA; 17799 MW; 0F67319F05EAC6E7 CRC64;

Query Match 92.2%; Score 130; DB 11; Length 157;
Best Local Similarity 93.3%; Pred. No. 2.6e-12;
Matches 28; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 AVPIAQKSEPHSLSSEALMRRAVSLVTDST 30
              |||||||||| |||:|||||||||||||||
Db         54 AVPIAQKSEPQSLSNEALMRRAVSLVTDST 83
```

RESULT 2
Q9AX95
ID Q9AX95 PRELIMINARY; PRT; 224 AA.
AC Q9AX95;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE P0501G01.22 protein.
GN P0501G01.22.
OS Oryza sativa (Rice).

```
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=4530;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Nipponbare;
RA   Sasaki T., Matsumoto T., Yamamoto K.;
RT   "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
RT   clone:P0501G01.";
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AP002819; BAB21093.1; -.
DR   Gramene; Q9AX95; -.
DR   InterPro; IPR001005; Myb_DNA_binding.
DR   PROSITE; PS00037; MYB_1; 1.
SQ   SEQUENCE   224 AA;  23798 MW;  FEF94A53AE500A92 CRC64;

  Query Match             39.7%;  Score 56;  DB 10;  Length 224;
  Best Local Similarity   37.0%;  Pred. No. 1.4;
  Matches   10;  Conservative    7;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          3 PIAQKSEPHSLSSEALMRRAVSLVTDS 29
              |:| :|:||: :: |:  |     |||
Db         93 PVATESQPHNTAARAVTGRVADASTDS 119


RESULT 3
Q18054
ID   Q18054      PRELIMINARY;      PRT;   313 AA.
AC   Q18054;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical 36.2 kDa protein.
GN   C17C3.7.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RX   MEDLINE=99069613; PubMed=9851916;
RA   None;
RT   "Genome sequence of the nematode C. elegans: a platform for
RT   investigating biology. The C. elegans Sequencing Consortium.";
RL   Science 282:2012-2018(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Du Z.;
RT   "The sequence of C. elegans cosmid C17C3.";
RL   Submitted (DEC-1995) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Waterston R.;
RT   "Direct Submission.";
RL   Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
CC       TRANSCRIPTION FACTORS.
DR   EMBL; U41279; AAK31421.1; -.
DR   WormPep; C17C3.7; CE04027.
DR   InterPro; IPR001092; HLH_basic.
DR   Pfam; PF00010; HLH; 1.
DR   SMART; SM00353; HLH; 1.
DR   PROSITE; PS50888; HLH_2; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   313 AA;  36167 MW;  DCCE0B2DAC63DF99 CRC64;

  Query Match             37.6%;  Score 53;  DB 5;  Length 313;
  Best Local Similarity   50.0%;  Pred. No. 6.1;
  Matches   12;  Conservative    3;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          4 IAQKSEPHSLSSEALMRRAVSLVT 27
              | ||||   :| | :: | | |||
Db         88 IVQKSEEEHISQEVVLFRIVKLVT 111


RESULT 4
Q18056
ID   Q18056      PRELIMINARY;      PRT;   321 AA.
AC   Q18056;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical 37.1 kDa protein.
GN   C17C3.10.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RX   MEDLINE=99069613; PubMed=9851916;
RA   None;
RT   "Genome sequence of the nematode C. elegans: a platform for
RT   investigating biology. The C. elegans Sequencing Consortium.";
RL   Science 282:2012-2018(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Du Z.;
RT   "The sequence of C. elegans cosmid C17C3.";
RL   Submitted (DEC-1995) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Waterston R.;
RT   "Direct Submission.";
RL   Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
CC       TRANSCRIPTION FACTORS.
DR   EMBL; U41279; AAK31423.1; -.
DR   WormPep; C17C3.10; CE04030.
DR   InterPro; IPR001092; HLH_basic.
DR   Pfam; PF00010; HLH; 1.
DR   SMART; SM00353; HLH; 1.
DR   PROSITE; PS50888; HLH_2; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   321 AA;  37103 MW;  7F3B63AA7549A1CD CRC64;

  Query Match             37.6%;  Score 53;  DB 5;  Length 321;
  Best Local Similarity   50.0%;  Pred. No. 6.3;
  Matches   12;  Conservative    3;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          4 IAQKSEPHSLSSEALMRRAVSLVT 27
              | ||||   :| | :: | | |||
Db         96 IVQKSEEEHISQEVVLFRIVKLVT 119


RESULT 5
Q8DF79
ID   Q8DF79      PRELIMINARY;      PRT;   181 AA.
AC   Q8DF79;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Cytosine/adenosine deaminase.
GN   VV10342.
OS   Vibrio vulnificus.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC   Vibrionaceae; Vibrio.
OX   NCBI_TaxID=672;
RN   [1]
```

```
RP   SEQUENCE FROM N.A.
RC   STRAIN=CMCP6;
RA   Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
RA   Choy H.E.;
RT   "Complete genome sequence of Vibrio vulnificus CMCP6.";
RL   Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AE016798; AAO08869.1; -.
KW   Complete proteome.
SQ   SEQUENCE   181 AA;  20108 MW;  E6FE855E85FF56AD CRC64;

  Query Match             36.9%;  Score 52;  DB 16;  Length 181;
  Best Local Similarity   45.5%;  Pred. No. 4.8;
  Matches   10;  Conservative    4;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          4 IAQKSEPHSLSSEALMRRAVSL 25
              :|:  :| ||  |  ||||: |
Db          1 MAETHQPFSLQDEVFMRRAIEL 22


RESULT 6
Q8DD72
ID   Q8DD72      PRELIMINARY;      PRT;   207 AA.
AC   Q8DD72;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Putative threonine efflux protein.
GN   VV11138.
OS   Vibrio vulnificus.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC   Vibrionaceae; Vibrio.
OX   NCBI_TaxID=672;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CMCP6;
RA   Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
RA   Choy H.E.;
RT   "Complete genome sequence of Vibrio vulnificus CMCP6.";
RL   Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AE016800; AAO09612.1; -.
KW   Complete proteome.
SQ   SEQUENCE   207 AA;  21859 MW;  ADA3B5E7F57E14E2 CRC64;

  Query Match             36.9%;  Score 52;  DB 16;  Length 207;
  Best Local Similarity   50.0%;  Pred. No. 5.6;
  Matches   10;  Conservative    6;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          4 IAQKSEPHSLSSEALMRRAV 23
              :|  |: |::|| ||:|:||
Db         99 LATASQQHAISSTALLRKAV 118


RESULT 7
P73661
ID   P73661      PRELIMINARY;      PRT;   630 AA.
AC   P73661;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical protein slr1888.
GN   SLR1888.
OS   Synechocystis sp. (strain PCC 6803).
OC   Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX   NCBI_TaxID=1148;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97061201; PubMed=8905231;
RA   Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA   Miyajima N., Hirosawa M., Sugiura M., Sasamoto S., Kimura T.,
RA   Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA   Shimpo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA   Tabata S.;
RT   "Sequence analysis of the genome of the unicellular cyanobacterium
RT   Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT   entire genome and assignment of potential protein-coding regions.";
RL   DNA Res. 3:109-136(1996).
DR   EMBL; D90908; BAA17706.1; -.
DR   InterPro; IPR003702; ActCoA_hydro.
DR   InterPro; IPR000182; GCN5acetyltransf.
DR   Pfam; PF02550; AcetylCoA_hydro; 1.
DR   Pfam; PF00583; Acetyltransf; 1.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   630 AA;  70937 MW;  457F6814AFDC7BBD CRC64;

  Query Match             36.9%;  Score 52;  DB 16;  Length 630;
  Best Local Similarity   40.6%;  Pred. No. 19;
  Matches   13;  Conservative    6;  Mismatches    7;  Indels    6;  Gaps    1;

Qy          3 PIAQKSEP------HSLSSEALMRRAVSLVTD 28
              |:    ||      :||| |:|:|| :|: ||
Db        479 PVKFNDEPNLKNFFYSLSDESLVRRFMSVRTD 510


RESULT 8
Q8AVG0
ID   Q8AVG0      PRELIMINARY;      PRT;   862 AA.
AC   Q8AVG0;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Similar to active BCR-related gene.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RA   Klein S., Strausberg R.;
RL   Submitted (JAN-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC042307; AAH42307.1; -.
SQ   SEQUENCE   862 AA;  98616 MW;  EC2AC4DB8F1E49FA CRC64;

  Query Match             36.9%;  Score 52;  DB 13;  Length 862;
  Best Local Similarity   39.3%;  Pred. No. 27;
  Matches   11;  Conservative    6;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          3 PIAQKSEPHSLSSEALMRRAVSLVTDST 30
              |   | :| |:: |||: : :  || ||
Db        222 PKDSKEQPQSVTMEALLYKPIDRVTRST 249


RESULT 9
Q9ZIG9
ID   Q9ZIG9      PRELIMINARY;      PRT;   196 AA.
AC   Q9ZIG9;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Putative ECF sigma factor X (RNA polymerase sigma factor).
GN   SIGX.
OS   Pseudomonas aeruginosa.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC   Pseudomonadaceae; Pseudomonas.
OX   NCBI_TaxID=287;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PAO1;
RX   MEDLINE=99369842; PubMed=10438740;
RA   Brinkman F.S., Schoofs G., Hancock R.E., De Mot R.;
RT   "Influence of a putative ECF sigma factor on expression of the major
RT   outer membrane protein, OprF, in Pseudomonas aeruginosa and
RT   Pseudomonas fluorescens.";
RL   J. Bacteriol. 181:4746-4754(1999).
DR   EMBL; AF027290; AAD11567.1; -.
```

```
SQ   SEQUENCE   196 AA;  23454 MW;  86CC12A4D48D4CF2 CRC64;

  Query Match             36.2%;  Score 51;  DB 2;  Length 196;
  Best Local Similarity   45.5%;  Pred. No. 7.5;
  Matches   10;  Conservative    5;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          3 PIAQKSEPHSLSSEALMRRAVS 24
              |::|: :|  || | |: || |
Db          6 PLSQRYDPRQLSDEELVERAHS 27


RESULT 10
Q8DC68
ID   Q8DC68      PRELIMINARY;      PRT;   928 AA.
AC   Q8DC68;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Signal transduction histidine kinase.
GN   VV11573.
OS   Vibrio vulnificus.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC   Vibrionaceae; Vibrio.
OX   NCBI_TaxID=672;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CMCP6;
RA   Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
RA   Choy H.E.;
RT   "Complete genome sequence of Vibrio vulnificus CMCP6.";
RL   Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AE016802; AAO09997.1; -.
KW   Kinase; Complete proteome.
SQ   SEQUENCE   928 AA;  103351 MW;  A2471C227B22C229 CRC64;

  Query Match             35.8%;  Score 50.5;  DB 16;  Length 928;
  Best Local Similarity   60.9%;  Pred. No. 50;
  Matches   14;  Conservative    1;  Mismatches    7;  Indels    1;  Gaps    1;

Qy          3 PIAQKSEPHSLS-SEALMRRAVS 24
              |||  |||| || |   :|| ||
Db         51 PIAIASEPHLLSESREAVRRLVS 73


RESULT 11
Q9RUK4
ID   Q9RUK4      PRELIMINARY;      PRT;   256 AA.
AC   Q9RUK4;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Hypothetical protein DR1380.
GN   DR1380.
OS   Deinococcus radiodurans.
OC   Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;
OC   Deinococcaceae; Deinococcus.
OX   NCBI_TaxID=1299;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=R1;
RX   MEDLINE=20036896; PubMed=10567266;
RA   White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA   Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA   Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA   Vamathevan J.J., Lam P., McDonald L., Utterback T., Zalewski C.,
RA   Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA   Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA   Fraser C.M.;
RT   "Genome sequence of the radioresistant bacterium Deinococcus
RT   radiodurans R1.";
RL   Science 286:1571-1577(1999).
DR   EMBL; AE001984; AAF10959.1; -.
DR   TIGR; DR1380; -.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   256 AA;  27643 MW;  E2A9A63C559077E3 CRC64;

  Query Match             35.5%;  Score 50;  DB 16;  Length 256;
  Best Local Similarity   38.1%;  Pred. No. 15;
  Matches    8;  Conservative    6;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          2 VPIAQKSEPHSLSSEALMRRA 22
              :|:| : :|| |:   |:| |
Db        188 IPLALRQDPHRLAQSGLVRAA 208


RESULT 12
Q9SZJ1
ID   Q9SZJ1      PRELIMINARY;      PRT;   720 AA.
AC   Q9SZJ1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical 79.2 kDa protein.
GN   F20D10.10 OR AT4G37890.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Bevan M., Wedler H., Kutzner M., Wambutt R., Bancroft I., Mewes H.W.,
RA   Mayer K.F.X., Schueller C.;
RL   Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   EU Arabidopsis sequencing project;
RL   Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Wedler H., Kutzner M., Wambutt R., Mewes H.W., Lemcke K.,
RA   Mayer K.F.X.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RA   EU Arabidopsis sequencing project;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
DR   EMBL; AL035538; CAB37529.1; -.
DR   EMBL; AL161592; CAB80454.1; -.
DR   InterPro; IPR002035; VWF_A.
DR   InterPro; IPR001841; Znf_ring.
DR   Pfam; PF00092; vwa; 1.
DR   Pfam; PF00097; zf-C3HC4; 1.
DR   SMART; SM00184; RING; 1.
DR   SMART; SM00327; VWA; 1.
DR   PROSITE; PS50089; ZF_RING_2; 1.
KW   Hypothetical protein; Metal-binding; Zinc; Zinc-finger.
SQ   SEQUENCE   720 AA;  79166 MW;  8E910981F41163EA CRC64;

  Query Match             35.5%;  Score 50;  DB 10;  Length 720;
  Best Local Similarity   33.3%;  Pred. No. 45;
  Matches   12;  Conservative    7;  Mismatches    7;  Indels   10;  Gaps    1;

Qy          3 PIAQKSEP----------HSLSSEALMRRAVSLVTD 28
              |: |||||            |:  |:||: :: |:|
Db        676 PVVQKSEPLTPTSAWRAAERLAKVAIMRKHMNRVSD 711


RESULT 13
Q8GYZ4
ID   Q8GYZ4      PRELIMINARY;      PRT;   739 AA.
AC   Q8GYZ4;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
```

```
DE   Hypothetical protein.
GN   AT4G37890/F20D10_10.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RA   Seki M., Iida K., Satou M., Sakurai T., Akiyama K., Ishida J.,
RA   Nakajima M., Enju A., Kamiya A., Narusaka M., Carninci P., Kawai J.,
RA   Hayashizaki Y., Shinozaki K.;
RT   "Arabidopsis thaliana full-length cDNA.";
RL   Submitted (NOV-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AK117301; BAC41972.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   739 AA;  81406 MW;  6F6949CACB24AE1C CRC64;

  Query Match            35.5%;  Score 50;  DB 10;  Length 739;
  Best Local Similarity  33.3%;  Pred. No. 47;
  Matches   12;  Conservative    7;  Mismatches    7;  Indels   10;  Gaps    1;

Qy          3 PIAQKSEP----------HSLSSEALMRRAVSLVTD 28
              |: ||||||            |:  |:||: :: |:|
Db        695 PVVQKSEPLTPTSAWRAAERLAKVAIMRKHMNRVSD 730


RESULT 14
Q9BYT9
ID   Q9BYT9      PRELIMINARY;      PRT;   981 AA.
AC   Q9BYT9;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Hypothetical protein.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Rosier M.F., Toselli E., Segurens-Soury B., Auffray C., Devignes M.D.;
RT   "Predominant brain expression and full-length characterization of a
RT   novel human 6.6-Kb transcript mapping at 11p14 in the telomeric part
RT   of WAGR locus.";
RL   Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Rosier M.F.;
RL   Thesis (1996), Department of Human Genetics, University Paris VII,
RL   Paris, France.
DR   EMBL; AJ300461; CAC32454.1; -.
DR   Genew; HGNC:14004; C11orf25.
KW   Hypothetical protein.
SQ   SEQUENCE   981 AA;  114654 MW;  15A3276420912393 CRC64;

  Query Match            35.5%;  Score 50;  DB 4;  Length 981;
  Best Local Similarity  47.8%;  Pred. No. 64;
  Matches   11;  Conservative    3;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          3 PIAQKSEPHSLSSEALMRRAVSL 25
              ||  | |||  ||: : |  ||:
Db        529 PITGKPEPHQPSSDKVTRLLVSV 551


RESULT 15
Q8EUZ4
ID   Q8EUZ4      PRELIMINARY;      PRT;  1003 AA.
AC   Q8EUZ4;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Conserved hypothetical protein.
GN   MYPE7730.
OS   Mycoplasma penetrans.
OC   Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX   NCBI_TaxID=28227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=HF-2;
RX   MEDLINE=22354719; PubMed=12466555;
RA   Sasaki Y., Ishikawa J., Yamashita A., Oshima K., Kenri T., Furuya K.,
RA   Yoshino C., Horino A., Shiba T., Sasaki T., Hattori M.;
RT   "The complete genomic sequence of Mycoplasma penetrans, an
RT   intracellular bacterial pathogen in humans.";
RL   Nucleic Acids Res. 30:5293-5300(2002).
DR   EMBL; AP004173; BAC44567.1; -.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   1003 AA;  109261 MW;  0C2AFF329C1D4F38 CRC64;

  Query Match            35.5%;  Score 50;  DB 16;  Length 1003;
  Best Local Similarity  37.5%;  Pred. No. 65;
  Matches    9;  Conservative    7;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          4 IAQKSEPHSLSSEALMRRAVSLVT 27
              : :: ||: :| | | | | |::|
Db        515 VVKRLEPNQISPELLKRYATSIIT 538


Search completed: October  2, 2003, 09:39:15
Job time : 105 secs
```

GenCore version 5.1.6


OM protein - protein search, using sw model

Run on: October 2, 2003, 09:35:28 ; Search time 25 Seconds
(without alignments)
50.773 Million cell updates/sec

Title: US-09-939-293A-19_COPY_56_85
Perfect score: 141
Sequence: 1 AVPIAQKSEPHSLSSEALMRRAVSLVTDST 30

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 141 | 100.0 | 239 | 3 | US-09-479-309-2 | Sequence 2, Appli |
| 2 | 141 | 100.0 | 239 | 4 | US-09-627-393-2 | Sequence 2, Appli |
| 3 | 51 | 36.2 | 198 | 4 | US-09-252-991A-30093 | Sequence 30093, A |
| 4 | 44.5 | 31.6 | 899 | 3 | US-09-413-814-5 | Sequence 5, Appli |
| 5 | 44 | 31.2 | 526 | 3 | US-09-342-648-8 | Sequence 8, Appli |
| 6 | 44 | 31.2 | 799 | 4 | US-09-165-396-4 | Sequence 4, Appli |
| 7 | 43.5 | 30.9 | 147 | 4 | US-09-252-991A-24435 | Sequence 24435, A |
| 8 | 43 | 30.5 | 249 | 2 | US-08-632-514C-11 | Sequence 11, Appl |
| 9 | 43 | 30.5 | 249 | 3 | US-09-188-177-11 | Sequence 11, Appl |
| 10 | 43 | 30.5 | 327 | 4 | US-09-252-991A-28715 | Sequence 28715, A |
| 11 | 43 | 30.5 | 369 | 3 | US-09-342-648-4 | Sequence 4, Appli |
| 12 | 43 | 30.5 | 570 | 3 | US-08-826-246-2 | Sequence 2, Appli |
| 13 | 43 | 30.5 | 570 | 3 | US-08-944-495-2 | Sequence 2, Appli |
| 14 | 43 | 30.5 | 570 | 3 | US-09-126-640-7 | Sequence 7, Appli |
| 15 | 43 | 30.5 | 570 | 3 | US-08-925-588-2 | Sequence 2, Appli |
| 16 | 43 | 30.5 | 570 | 4 | US-09-288-292A-7 | Sequence 7, Appli |
| 17 | 43 | 30.5 | 570 | 4 | US-09-372-044-2 | Sequence 2, Appli |
| 18 | 43 | 30.5 | 570 | 4 | US-08-825-486-2 | Sequence 2, Appli |
| 19 | 42 | 29.8 | 212 | 4 | US-09-252-991A-32491 | Sequence 32491, A |
| 20 | 42 | 29.8 | 570 | 3 | US-08-747-221B-54 | Sequence 54, Appl |
| 21 | 42 | 29.8 | 570 | 3 | US-09-005-051-54 | Sequence 54, Appl |
| 22 | 42 | 29.8 | 596 | 3 | US-08-747-221B-25 | Sequence 25, Appl |
| 23 | 42 | 29.8 | 596 | 3 | US-09-005-051-25 | Sequence 25, Appl |
| 24 | 42 | 29.8 | 640 | 4 | US-09-252-991A-23007 | Sequence 23007, A |
| 25 | 42 | 29.8 | 755 | 3 | US-09-342-648-2 | Sequence 2, Appli |
| 26 | 42 | 29.8 | 1014 | 4 | US-09-252-991A-17583 | Sequence 17583, A |
| 27 | 42 | 29.8 | 1201 | 3 | US-09-098-901-2 | Sequence 2, Appli |
| 28 | 41 | 29.1 | 161 | 4 | US-09-252-991A-31686 | Sequence 31686, A |
| 29 | 41 | 29.1 | 262 | 4 | US-09-107-532A-5791 | Sequence 5791, Ap |
| 30 | 41 | 29.1 | 383 | 4 | US-09-252-991A-29621 | Sequence 29621, A |
| 31 | 41 | 29.1 | 460 | 4 | US-09-198-452A-7 | Sequence 7, Appli |
| 32 | 41 | 29.1 | 622 | 3 | US-09-342-648-6 | Sequence 6, Appli |
| 33 | 41 | 29.1 | 1027 | 4 | US-09-252-991A-23210 | Sequence 23210, A |
| 34 | 41 | 29.1 | 3079 | 5 | PCT-US94-00198-4 | Sequence 4, Appli |
| 35 | 40.5 | 28.7 | 189 | 2 | US-08-861-269-7 | Sequence 7, Appli |
| 36 | 40.5 | 28.7 | 189 | 2 | US-09-134-596-7 | Sequence 7, Appli |
| 37 | 40.5 | 28.7 | 189 | 3 | US-09-293-273-7 | Sequence 7, Appli |
| 38 | 40.5 | 28.7 | 441 | 4 | US-09-198-452A-1124 | Sequence 1124, Ap |
| 39 | 40 | 28.4 | 144 | 4 | US-09-252-991A-31261 | Sequence 31261, A |
| 40 | 40 | 28.4 | 149 | 4 | US-09-347-650-8 | Sequence 8, Appli |
| 41 | 40 | 28.4 | 175 | 4 | US-09-252-991A-18834 | Sequence 18834, A |
| 42 | 40 | 28.4 | 179 | 4 | US-09-252-991A-28433 | Sequence 28433, A |
| 43 | 40 | 28.4 | 242 | 4 | US-09-198-452A-1006 | Sequence 1006, Ap |
| 44 | 40 | 28.4 | 323 | 4 | US-09-252-991A-29849 | Sequence 29849, A |
| 45 | 40 | 28.4 | 330 | 4 | US-09-252-991A-26360 | Sequence 26360, A |

ALIGNMENTS

RESULT 1
US-09-479-309-2
; Sequence 2, Application US/09479309
; Patent No. 6110691
; GENERAL INFORMATION:
; APPLICANT: Wang, Xiaodong
; APPLICANT: Du, Chunying
; TITLE OF INVENTION: Activators of Caspases
; FILE REFERENCE: UTSD0630
; CURRENT APPLICATION NUMBER: US/09/479,309
; CURRENT FILING DATE: 2000-01-06
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 239
; TYPE: PRT
; ORGANISM: human
US-09-479-309-2

Query Match 100.0%; Score 141; DB 3; Length 239;
Best Local Similarity 100.0%; Pred. No. 1.2e-15;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AVPIAQKSEPHSLSSEALMRRAVSLVTDST 30
||||||||||||||||||||||||||||||
Db 56 AVPIAQKSEPHSLSSEALMRRAVSLVTDST 85

RESULT 2
US-09-627-393-2
; Sequence 2, Application US/09627393
; Patent No. 6534267
; GENERAL INFORMATION:
; APPLICANT: Wang, Xiaodong
; APPLICANT: Du, Chunying
; TITLE OF INVENTION: Activators of Caspases
; FILE REFERENCE: UTSD0630
; CURRENT APPLICATION NUMBER: US/09/627,393
; CURRENT FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: 09/479,309
; PRIOR FILING DATE: 2000-01-06
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 239
; TYPE: PRT
; ORGANISM: human
US-09-627-393-2

Query Match 100.0%; Score 141; DB 4; Length 239;

```
  Best Local Similarity   100.0%;  Pred. No. 1.2e-15;
  Matches   30;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AVPIAQKSEPHSLSSEALMRRAVSLVTDST 30
              ||||||||||||||||||||||||||||||
Db         56 AVPIAQKSEPHSLSSEALMRRAVSLVTDST 85


RESULT 3
US-09-252-991A-30093
; Sequence 30093, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 30093
;   LENGTH: 198
;   TYPE: PRT
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30093

  Query Match             36.2%;  Score 51;  DB 4;  Length 198;
  Best Local Similarity   45.5%;  Pred. No. 1;
  Matches   10;  Conservative    5;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          3 PIAQKSEPHSLSSEALMRRAVS 24
              |::|: :|  || | |: || |
Db          8 PLSQRYDPRQLSDEELVERAHS 29


RESULT 4
US-09-413-814-5
; Sequence 5, Application US/09413814
; Patent No. 6225064
; GENERAL INFORMATION:
;  APPLICANT: Gesellschaft fuer Biotechnologische Forschung mbH
;  APPLICANT: Bristol-Myers Squibb, Co.
;  APPLICANT: Beyer, Stefan
;  APPLICANT: Bloecker, Helmut
;  APPLICANT: Brandt, Petra
;  APPLICANT: Cino, Paul M
;  APPLICANT: Dougherty, Brian A
;  APPLICANT: Goldberg, Steven L
;  APPLICANT: Hofle, Gerhard
;  APPLICANT: Mueller, Joachim
;  APPLICANT: Reichenbach, Hans
;  TITLE OF INVENTION: DNA sequences for enzymatic synthesis of polyketide or
;  TITLE OF INVENTION: heteropolyketide compounds
;  FILE REFERENCE: PCT/US 99/23535
;  CURRENT APPLICATION NUMBER: US/09/413,814
;  CURRENT FILING DATE: 1999-10-07
;  EARLIER APPLICATION NUMBER: DE 198 46 493.2
;  EARLIER FILING DATE: 1998-10-09
;  NUMBER OF SEQ ID NOS: 107
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
;   LENGTH: 899
;   TYPE: PRT
;   ORGANISM: Sorangium cellulosum
US-09-413-814-5

  Query Match             31.6%;  Score 44.5;  DB 3;  Length 899;
  Best Local Similarity   33.3%;  Pred. No. 88;
  Matches   10;  Conservative    9;  Mismatches    8;  Indels    3;  Gaps    1;

Qy          4 IAQKSEPHSLSSE---ALMRRAVSLVTDST 30
              ::| :| | :| |   ||: :  ::| |:|
Db        665 LSQSAEVHPISEELDRALLHKLSAVVDDAT 694


RESULT 5
US-09-342-648-8
; Sequence 8, Application US/09342648
; Patent No. 6248584
; GENERAL INFORMATION:
;  APPLICANT: Cahoon, Rebecca E.
;  APPLICANT: Odell, Joan
;  APPLICANT: Rafalski, Antoni
;  TITLE OF INVENTION: Transcription Coactivators
;  FILE REFERENCE: BB-1169-B
;  CURRENT APPLICATION NUMBER: US/09/342,648
;  CURRENT FILING DATE: 1999-06-29
;  EARLIER APPLICATION NUMBER: 60/092,659
;  EARLIER FILING DATE: July 13, 1998
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: Microsoft Office 97
; SEQ ID NO 8
;   LENGTH: 526
;   TYPE: PRT
;   ORGANISM: Triticum aestivum
US-09-342-648-8

  Query Match             31.2%;  Score 44;  DB 3;  Length 526;
  Best Local Similarity   52.9%;  Pred. No. 53;
  Matches    9;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          7 KSEPHSLSSEALMRRAV 23
              | ||:|  : |||||
Db        163 KDEPYSSEAIALMRRMI 179


RESULT 6
US-09-165-396-4
; Sequence 4, Application US/09165396
; Patent No. 6441134
; GENERAL INFORMATION:
;  APPLICANT: BECKER, JEFFREY W.
;  APPLICANT: LUBKOWITZ, MARK A.
;  TITLE OF INVENTION: ISOLATED CANDIDA ALBICANS OLIGOPEPTIDE TRANSPORTER GENE
;  FILE REFERENCE: 372.6520P
;  CURRENT APPLICATION NUMBER: US/09/165,396
;  CURRENT FILING DATE: 1998-10-02
;  EARLIER APPLICATION NUMBER: PCT/US98/02332
;  EARLIER FILING DATE: 1998-02-06
;  EARLIER APPLICATION NUMBER: 60/037,859
;  EARLIER FILING DATE: 1997-02-07
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
;   LENGTH: 799
;   TYPE: PRT
;   ORGANISM: S. cerevisiae
US-09-165-396-4

  Query Match             31.2%;  Score 44;  DB 4;  Length 799;
  Best Local Similarity   29.5%;  Pred. No. 91;
  Matches   13;  Conservative    5;  Mismatches   10;  Indels   16;  Gaps    1;

Qy          3 PIAQKSEPHSLSSEAL----------------MRRAVSLVTDST 30
              | :|| : ||:  | |                :| |||:  | |
Db         58 PESQKFDRHSIQEEGLVWKGDPTYLPNSPYPEVRSAVSIEDDPT 101


RESULT 7
US-09-252-991A-24435
; Sequence 24435, Application US/09252991A
; Patent No. 6551795
```

```
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 24435
;   LENGTH: 147
;   TYPE: PRT
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24435

  Query Match             30.9%;  Score 43.5;  DB 4;  Length 147;
  Best Local Similarity   34.5%;  Pred. No. 12;
  Matches   10;  Conservative   10;  Mismatches    8;  Indels    1;  Gaps    1;

Qy          1 AVPIAQKSEPHSLSSEALMRRAVSLVTDS 29
              |: :|::| |:|| : |:   :|:|  :|
Db          8 AIRLAKRSSPYSLKTRAI-SSSVALAMNS 35


RESULT 8
US-08-632-514C-11
; Sequence 11, Application US/08632514C
; Patent No. 5834234
;  GENERAL INFORMATION:
;    APPLICANT:  GALLO, Gregory J.
;    TITLE OF INVENTION:  APOPTOSIS ASSOCIATED PROTEIN BBK
;    NUMBER OF SEQUENCES:  29
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Hale and Dorr
;      STREET:  1455 Pennsylvania Avenue, N.W.
;      CITY:  Washington, D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/632,514C
;      FILING DATE:  29-MAY-1996
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  WIXON, Henry N.
;      REGISTRATION NUMBER:  32,073
;      REFERENCE/DOCKET NUMBER:  104322.188
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)942-8459
;      TELEFAX:  (202)942-8484
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  249 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-632-514C-11

  Query Match             30.5%;  Score 43;  DB 2;  Length 249;
  Best Local Similarity   34.6%;  Pred. No. 29;
  Matches    9;  Conservative    6;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          3 PIAQKSEPHSLSSEALMRRAVSLVTD 28
              |  :   | |   ||::|| |:|: :
Db        109 PSPELQGPPSTEKEAILRRLVALLEE 134


RESULT 9
US-09-188-177-11
; Sequence 11, Application US/09188177
; Patent No. 6057132
;  GENERAL INFORMATION:
;    APPLICANT:  GALLO, Gregory J.
;    TITLE OF INVENTION:  APOPTOSIS ASSOCIATED PROTEIN BBK
;    NUMBER OF SEQUENCES:  29
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Hale and Dorr
;      STREET:  1455 Pennsylvania Avenue, N.W.
;      CITY:  Washington, D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/188,177
;      FILING DATE:  2
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  WIXON, Henry N.
;      REGISTRATION NUMBER:  32,073
;      REFERENCE/DOCKET NUMBER:  104322.188
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)942-8459
;      TELEFAX:  (202)942-8484
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  249 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-188-177-11

  Query Match             30.5%;  Score 43;  DB 3;  Length 249;
  Best Local Similarity   34.6%;  Pred. No. 29;
  Matches    9;  Conservative    6;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          3 PIAQKSEPHSLSSEALMRRAVSLVTD 28
              |  :   | |   ||::|| |:|: :
Db        109 PSPELQGPPSTEKEAILRRLVALLEE 134


RESULT 10
US-09-252-991A-28715
; Sequence 28715, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 28715
;   LENGTH: 327
;   TYPE: PRT
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28715

  Query Match             30.5%;  Score 43;  DB 4;  Length 327;
```

```
  Best Local Similarity   32.1%;  Pred. No. 42;
  Matches    9;  Conservative    7;  Mismatches   12;  Indels    0;  Gaps    0;

Qy          2 VPIAQKSEPHSLSSEALMRRAVSLVTDS 29
              ||  ::  || |: : |  ||::   |:
Db         16 VPTGRRVLPHPLAGQRLRHRALAAHGDA 43


RESULT 11
US-09-342-648-4
; Sequence 4, Application US/09342648
; Patent No. 6248584
; GENERAL INFORMATION:
;  APPLICANT: Cahoon, Rebecca E.
;  APPLICANT: Odell, Joan
;  APPLICANT: Rafalski, Antoni
;  TITLE OF INVENTION: Transcription Coactivators
;  FILE REFERENCE: BB-1169-B
;  CURRENT APPLICATION NUMBER: US/09/342,648
;  CURRENT FILING DATE: 1999-06-29
;  EARLIER APPLICATION NUMBER: 60/092,659
;  EARLIER FILING DATE: July 13, 1998
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: Microsoft Office 97
; SEQ ID NO 4
;   LENGTH: 369
;   TYPE: PRT
;   ORGANISM: Oryza sativa
US-09-342-648-4

  Query Match             30.5%;  Score 43;  DB 3;  Length 369;
  Best Local Similarity   52.9%;  Pred. No. 49;
  Matches    9;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          7 KSEPHSLSSEALMRRAV 23
              | ||:|  : ||||| :
Db          6 KDEPYSNEAIALMRRRI 22


RESULT 12
US-08-826-246-2
; Sequence 2, Application US/08826246
; Patent No. 6048709
;  GENERAL INFORMATION:
;    APPLICANT:  Falb, Dean
;    TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR
;    TITLE OF INVENTION:  THE TREATMENT AND DIAGNOSIS OF
;    TITLE OF INVENTION:  CARDIOVASCULAR DISEASE
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PENNIE & EDMONDS LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10036-2711
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/826,246
;      FILING DATE:  28-MAR-1997
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/799,910
;      FILING DATE:  13-FEB-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/011,787
;      FILING DATE:  16-FEB-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Coruzzi, Laura A
;      REGISTRATION NUMBER:  30,742
;      REFERENCE/DOCKET NUMBER:  7853-078-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212)7909090
;      TELEFAX:  (212)8699741
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  570 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  protein
;    FRAGMENT TYPE:  internal
US-08-826-246-2

  Query Match             30.5%;  Score 43;  DB 3;  Length 570;
  Best Local Similarity   43.5%;  Pred. No. 86;
  Matches   10;  Conservative    4;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          7 KSEPHSLSSEALMRRAVSLVTDS 29
              || |  | :|||   | :|:| :
Db        451 KSHPEVLIAEALAANAGALITST 473


RESULT 13
US-08-944-495-2
; Sequence 2, Application US/08944495
; Patent No. 6087477
;  GENERAL INFORMATION:
;    APPLICANT:  Falb, Dean
;    TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR
;    TITLE OF INVENTION:  THE TREATMENT AND DIAGNOSIS OF
;    TITLE OF INVENTION:  CARDIOVASCULAR DISEASE
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PENNIE & EDMONDS LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10036-2711
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/944,495
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/799,910
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Coruzzi, Laura A
;      REGISTRATION NUMBER:  30,742
;      REFERENCE/DOCKET NUMBER:  7853-067-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212)7909090
;      TELEFAX:  (212)8699741
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  570 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  protein
;    FRAGMENT TYPE:  internal
US-08-944-495-2

  Query Match             30.5%;  Score 43;  DB 3;  Length 570;
```

```
  Best Local Similarity   43.5%;  Pred. No. 86;
  Matches   10;  Conservative    4;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          7 KSEPHSLSSEALMRRAVSLVTDS 29
              || |  | :|||   | :|:| :
Db        451 KSHPEVLIAEALAANAGALITST 473


RESULT 14
US-09-126-640-7
; Sequence 7, Application US/09126640A
; Patent No. 6099823
; GENERAL INFORMATION:
;  APPLICANT: FALB, Dean A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
;  TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF CARDIOVASCULAR DISEASE
;  FILE REFERENCE: 7853-126
;  CURRENT APPLICATION NUMBER: US/09/126,640A
;  CURRENT FILING DATE: 1998-07-30
;  EARLIER APPLICATION NUMBER: 08/870,434
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 08/799,910
;  EARLIER FILING DATE: 1997-02-13
;  EARLIER APPLICATION NUMBER: 60/011,787
;  EARLIER FILING DATE: 1996-02-16
;  NUMBER OF SEQ ID NOS: 44
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7
;   LENGTH: 570
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-126-640-7

  Query Match            30.5%;  Score 43;  DB 3;  Length 570;
  Best Local Similarity   43.5%;  Pred. No. 86;
  Matches   10;  Conservative    4;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          7 KSEPHSLSSEALMRRAVSLVTDS 29
              || |  | :|||   | :|:| :
Db        451 KSHPEVLIAEALAANAGALITST 473


RESULT 15
US-08-925-588-2
; Sequence 2, Application US/08925588
; Patent No. 6221628
;   GENERAL INFORMATION:
;        APPLICANT: Falb, Dean
;        TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
;                            THE TREATMENT AND DIAGNOSIS OF
;                            CARDIOVASCULAR DISEASE
;        NUMBER OF SEQUENCES: 44
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: PENNIE & EDMONDS LLP
;             STREET: 1155 Avenue of the Americas
;             CITY: New York
;             STATE: NY
;             COUNTRY: USA
;             ZIP: 10036-2711
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette
;             COMPUTER: IBM Compatible
;             OPERATING SYSTEM: DOS
;             SOFTWARE: FastSEQ Version 2.0
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/08/925,588
;             FILING DATE: 08-Sep-1997
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/799,910
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Coruzzi, Laura A
;             REGISTRATION NUMBER: 30,742
;             REFERENCE/DOCKET NUMBER: 7853-067-999
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (212)7909090
;             TELEFAX: (212)8699741
;             TELEX: 66141 PENNIE
;   INFORMATION FOR SEQ ID NO: 2:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 570 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: unknown
;             TOPOLOGY: unknown
;        MOLECULE TYPE: protein
;        FRAGMENT TYPE: internal
;        SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-08-925-588-2

  Query Match            30.5%;  Score 43;  DB 3;  Length 570;
  Best Local Similarity   43.5%;  Pred. No. 86;
  Matches   10;  Conservative    4;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          7 KSEPHSLSSEALMRRAVSLVTDS 29
              || |  | :|||   | :|:| :
Db        451 KSHPEVLIAEALAANAGALITST 473


Search completed: October  2, 2003, 09:40:40
Job time : 26 secs
```

GenCore version 5.1.6


OM protein - protein search, using sw model

Run on: October 2, 2003, 09:40:43 ; Search time 18 Seconds
(without alignments)
160.281 Million cell updates/sec

Title: US-09-939-293A-19_COPY_56_85
Perfect score: 141
Sequence: 1 AVPIAQKSEPHSLSSEALMRRAVSLVTDST 30

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 6280

Minimum DB seq length: 0
Maximum DB seq length: 30

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 35 | 24.8 | 20 | 2 | T48881 | leader peptide [im |
| 2 | 35 | 24.8 | 23 | 2 | I51922 | cystic fibrosis tr |
| 3 | 34 | 24.1 | 15 | 1 | LFTWL | leu leader peptide |
| 4 | 32 | 22.7 | 27 | 2 | S00347 | triacylglycerol li |
| 5 | 29 | 20.6 | 20 | 2 | A60897 | class I histocompa |
| 6 | 29 | 20.6 | 28 | 2 | S04341 | cytochrome P450 PB |
| 7 | 27 | 19.1 | 18 | 2 | G42753 | interferon alpha ( |
| 8 | 27 | 19.1 | 20 | 2 | A60728 | cytochrome P450 3A |
| 9 | 27 | 19.1 | 22 | 2 | C48186 | ATPase R1 subunit |
| 10 | 27 | 19.1 | 23 | 2 | A48186 | ATP synthase beta- |
| 11 | 27 | 19.1 | 27 | 2 | A30323 | amyloid protein AL |
| 12 | 26 | 18.4 | 24 | 2 | T29626 | hypothetical prote |
| 13 | 26 | 18.4 | 24 | 2 | A37825 | fibronectin recept |
| 14 | 26 | 18.4 | 27 | 2 | A43768 | Hu-like protein HB |
| 15 | 26 | 18.4 | 29 | 2 | A60558 | cytochrome P450 HL |
| 16 | 26 | 18.4 | 29 | 2 | S17432 | H+-transporting tw |
| 17 | 26 | 18.4 | 29 | 2 | S01614 | dystrophin - rat ( |
| 18 | 25 | 17.7 | 14 | 4 | S00843 | hypothetical prote |
| 19 | 25 | 17.7 | 15 | 2 | PU0025 | ubiquinol-cytochro |
| 20 | 25 | 17.7 | 22 | 2 | A37391 | sex pheromone inhi |
| 21 | 25 | 17.7 | 23 | 2 | T10123 | probable catalase |
| 22 | 25 | 17.7 | 24 | 2 | S55764 | cathepsin G (EC 3. |
| 23 | 25 | 17.7 | 24 | 2 | I39289 | ZF3 domain - human |
| 24 | 25 | 17.7 | 24 | 4 | S09363 | hypothetical MTCO1 |
| 25 | 25 | 17.7 | 25 | 2 | E42753 | interferon alpha ( |
| 26 | 25 | 17.7 | 27 | 2 | S00735 | probable membrane |
| 27 | 25 | 17.7 | 28 | 2 | S70894 | hypothetical prote |
| 28 | 25 | 17.7 | 28 | 2 | PL0005 | pepsin A (EC 3.4.2 |
| 29 | 25 | 17.7 | 29 | 2 | A61613 | ceratotoxin A - Me |
| 30 | 25 | 17.7 | 29 | 2 | B61613 | ceratotoxin B - Me |
| 31 | 25 | 17.7 | 30 | 2 | S07217 | ribosomal protein |
| 32 | 24.5 | 17.4 | 23 | 2 | A59480 | NADP phosphatase I |
| 33 | 24.5 | 17.4 | 25 | 2 | PC4445 | L-ascorbate peroxi |
| 34 | 24 | 17.0 | 12 | 1 | A43975 | locustamyotropin - |
| 35 | 24 | 17.0 | 17 | 2 | S71327 | hypothetical prote |
| 36 | 24 | 17.0 | 18 | 2 | S55002 | protein 1 - Legion |
| 37 | 24 | 17.0 | 20 | 2 | H28949 | ribosomal protein |
| 38 | 24 | 17.0 | 21 | 2 | A35646 | mast cell proteina |
| 39 | 24 | 17.0 | 21 | 2 | A59325 | probable bacteriop |
| 40 | 24 | 17.0 | 25 | 2 | PQ0369 | L protein - rabies |
| 41 | 24 | 17.0 | 25 | 2 | A24807 | cytotoxic T-lympho |
| 42 | 24 | 17.0 | 25 | 2 | D20554 | hemocyanin subunit |
| 43 | 24 | 17.0 | 26 | 2 | B24743 | prolactin, 24K - M |
| 44 | 24 | 17.0 | 26 | 2 | H42753 | interferon alpha ( |
| 45 | 24 | 17.0 | 27 | 2 | T13836 | NADH2 dehydrogenas |

ALIGNMENTS

RESULT 1
T48881
leader peptide [imported] - Vibrio sp.
C;Species: Vibrio sp.
C;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
C;Accession: T48881
R;Xu, Y.; Zhang, Y.; Liang, Z.Y.; Van de Casteele, M.; Legrain, C.; Glansdorff, N.
Microbiology 144, 1435-1441, 1998
A;Title: Aspartate carbamoyltransferase from a psychrophilic deep-sea bacterium, Vibr
A;Reference number: Z24845
A;Accession: T48881
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-20 <XUY>
A;Cross-references: EMBL:Y09786; PIDN:CAA70922.1
A;Experimental source: strain 2693

Query Match 24.8%; Score 35; DB 2; Length 20;
Best Local Similarity 53.3%; Pred. No. 58;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

```
Qy          6 QKSEPHSLSSEALMR 20
              |:: | ||||  |:|
Db          2 QRAAPSSLSSFKLVR 16
```

RESULT 2
I51922
cystic fibrosis transmembrane conductance regulator - rabbit (fragment)
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 04-Sep-1997 #sequence_revision 07-Nov-1997 #text_change 20-Aug-1999
C;Accession: I51922
R;McGrath, S.A.; Basu, A.; Zeitlin, P.L.
Am. J. Respir. Cell Mol. Biol. 8, 201-208, 1993
A;Title: Cystic fibrosis gene and protein expression during fetal lung development.
A;Reference number: I51922; MUID:93152187; PMID:7678968
A;Accession: I51922
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-23 <MCG>
A;Cross-references: GB:S54552; NID:g265093; PIDN:AAB25301.1; PID:g265094
C;Superfamily: cystic fibrosis transmembrane conductance regulator; ATP-binding casse

Query Match 24.8%; Score 35; DB 2; Length 23;
Best Local Similarity 53.3%; Pred. No. 68;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

```
Qy         15 SEALMRRAVSLVTDS 29
              |:| : | :||| ||
Db          5 SDASIERRLSLVPDS 19
```

```
RESULT 3
LFTWL
leu leader peptide - Thermus aquaticus
C;Species: Thermus aquaticus
C;Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 16-Jul-1999
C;Accession: S00901
R;Croft, J.E.; Love, D.R.; Bergquist, P.L.
Mol. Gen. Genet. 210, 490-497, 1987
A;Title: Expression of leucine genes from an extremely thermophilic bacterium in Escheri
A;Reference number: S00901; MUID:88121725; PMID:3323845
A;Accession: S00901
A;Molecule type: DNA
A;Residues: 1-15 <CRO>
A;Cross-references: EMBL:X06604; NID:g48244; PIDN:CAA29823.1; PID:g48245
A;Note: the source is designated as Thermus thermophilus
C;Superfamily: Thermus aquaticus leu leader peptide

  Query Match             24.1%;  Score 34;  DB 1;  Length 15;
  Best Local Similarity   80.0%;  Pred. No. 59;
  Matches    8;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         19 MRRAVSLVTD 28
              ||||| || |
Db          1 MRRAVILVLD 10


RESULT 4
S00347
triacylglycerol lipase (EC 3.1.1.3) - bovine (fragment)
N;Alternate names: hormone-sensitive lipase
C;Species: Bos primigenius taurus (cattle)
C;Date: 30-Sep-1989 #sequence_revision 30-Sep-1989 #text_change 30-Sep-1993
C;Accession: S00347
R;Garton, A.J.; Campbell, D.G.; Cohen, P.; Yeaman, S.J.
FEBS Lett. 229, 68-72, 1988
A;Title: Primary structure of the site on bovine hormone-sensitive lipase phosphorylated
A;Reference number: S00347; MUID:88152238; PMID:3345839
A;Accession: S00347
A;Molecule type: protein
A;Residues: 1-27 <GAR>
C;Comment: Activation of this enzyme involves phosphorylation of Ser-8 by cyclic AMP-dep
C;Superfamily: hormone-sensitive lipase
C;Keywords: carboxylic ester hydrolase; lipid degradation; phosphoprotein
F;8/Binding site: phosphate (Ser) (covalent) #status experimental
F;10/Binding site: phosphate (Ser) (covalent) #status predicted

  Query Match             22.7%;  Score 32;  DB 2;  Length 27;
  Best Local Similarity   44.0%;  Pred. No. 2.4e+02;
  Matches   11;  Conservative    3;  Mismatches    9;  Indels    2;  Gaps    1;

Qy          7 KSEP--HSLSSEALMRRAVSLVTDS 29
              |:||   |:|  || :    | |||
Db          1 KTEPMRRSVSEAALTQPEGPLGTDS 25


RESULT 5
A60897
class I histocompatibility antigen H-2K(d) alpha chain, alternate splice form - mouse (f
C;Species: Mus musculus (house mouse)
C;Date: 04-Nov-1994 #sequence_revision 04-Nov-1994 #text_change 07-May-1999
C;Accession: A60897
R;Abu-Hadid, M.M.; Fuji, H.; Sood, A.K.
Mol. Immunol. 25, 739-749, 1988
A;Title: Identification of an alternatively spliced K(d) and the Qa-6(d) mRNAs by using
A;Reference number: A60897; MUID:89039921; PMID:3141798
A;Accession: A60897
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-20 <ABU>

  Query Match             20.6%;  Score 29;  DB 2;  Length 20;
  Best Local Similarity   35.7%;  Pred. No. 4.9e+02;
  Matches    5;  Conservative    4;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 AVPIAQKSEPHSLS 14
              ::|     :||||:
Db          7 SLPDVMVHDPHSLA 20


RESULT 6
S04341
cytochrome P450 PBD-1 - dog (fragment)
N;Alternate names: cytochrome P450 3A
N;Contains: oxidoreductase (EC 1.-.-.-)
C;Species: Canis lupus familiaris (dog)
C;Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 05-Mar-1999
C;Accession: S04341
R;Ciaccio, P.J.; Halpert, J.R.
Arch. Biochem. Biophys. 271, 284-299, 1989
A;Title: Characterization of a phenobarbital-inducible dog liver cytochrome P450 stru
A;Reference number: S04341; MUID:89271912; PMID:2786372
A;Accession: S04341
A;Molecule type: protein
A;Residues: 1-28 <CIA>
C;Genetics:
A;Gene: CYP3A
C;Superfamily: human cytochrome P450 CYP3A5; cytochrome P450 homology
C;Keywords: electron transfer; endoplasmic reticulum; heme; monooxygenase; oxidoreduc

  Query Match             20.6%;  Score 29;  DB 2;  Length 28;
  Best Local Similarity   46.7%;  Pred. No. 7.3e+02;
  Matches    7;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy         12 SLSSEALMRRAVSLV 26
              | |:|  :  |:|||
Db          6 SFSTETWLLLAISLV 20


RESULT 7
G42753
interferon alpha (component i) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 09-Sep-1994 #sequence_revision 09-Sep-1994 #text_change 17-Mar-1999
C;Accession: G42753
R;Zoon, K.C.; Miller, D.; Bekisz, J.; zur Nedden, D.; Enterline, J.C.; Nguyen, N.Y.;
J. Biol. Chem. 267, 15210-15216, 1992
A;Title: Purification and characterization of multiple components of human lymphoblas
A;Reference number: A42753; MUID:92340576; PMID:1634550
A;Accession: G42753
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-18 <ZOO>

  Query Match             19.1%;  Score 27;  DB 2;  Length 18;
  Best Local Similarity   44.4%;  Pred. No. 8.9e+02;
  Matches    8;  Conservative    3;  Mismatches    3;  Indels    4;  Gaps    1;

Qy          9 EPHSLSSEALMRRAVSLV 26
              | ||| :    |||: |:
Db          5 ETHSLDN----RRALILL 18


RESULT 8
A60728
cytochrome P450 3A, troleandomycin-induced - sheep (fragment)
N;Contains: oxidoreductase (EC 1.-.-.-)
C;Species: Ovis sp. (sheep)
C;Date: 14-May-1993 #sequence_revision 14-May-1993 #text_change 05-Mar-1999
C;Accession: A60728
R;Pineau, T.; Galtier, P.; Bonfils, C.; Derancourt, J.; Maurel, P.
Biochem. Pharmacol. 39, 901-909, 1990
A;Title: Purification of a sheep liver cytochrome P-450 from the P450IIIA gene subfam
A;Reference number: A60728; MUID:90179800; PMID:2310415
A;Accession: A60728
A;Molecule type: protein
A;Residues: 1-20 <PIN>
```

```
C;Comment: This cytochrome P450 isozyme is a member of the P450IIIA family but is not fu
C;Genetics:
A;Gene: CYP3A
C;Superfamily: human cytochrome P450 CYP3A5; cytochrome P450 homology
C;Keywords: electron transfer; endoplasmic reticulum; heme; monooxygenase; oxidoreductas

  Query Match            19.1%;  Score 27;  DB 2;  Length 20;
  Best Local Similarity   46.7%;  Pred. No. 1e+03;
  Matches    7;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy         12 SLSSEALMRRAVSLV 26
              | | |  :  |:|||
Db          6 SFSKETWVLLAISLV 20


RESULT 9
C48186
ATPase R1 subunit - wood tobacco (fragment)
C;Species: Nicotiana sylvestris (wood tobacco)
C;Date: 16-Feb-1994 #sequence_revision 18-Nov-1994 #text_change 23-Feb-1997
C;Accession: C48186
R;De Paepe, R.; Forchioni, A.; Chetrit, P.; Vedel, F.
Proc. Natl. Acad. Sci. U.S.A. 90, 5934-5938, 1993
A;Title: Specific mitochondrial proteins in pollen: presence of an additional ATP syntha
A;Reference number: A48186; MUID:93317598; PMID:8327463
A;Accession: C48186
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-22 <DE1>
A;Experimental source: pollen
A;Note: sequence extracted from NCBI backbone (NCBIP:134869)

  Query Match            19.1%;  Score 27;  DB 2;  Length 22;
  Best Local Similarity   50.0%;  Pred. No. 1.1e+03;
  Matches    6;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 AVPIAQKSEPHS 12
              | |  ||  | |
Db          8 ATPAPQKKTPGS 19


RESULT 10
A48186
ATP synthase beta-1 chain - wood tobacco (fragment)
C;Species: Nicotiana sylvestris (wood tobacco)
C;Date: 16-Feb-1994 #sequence_revision 18-Nov-1994 #text_change 23-Mar-1995
C;Accession: A48186
R;De Paepe, R.; Forchioni, A.; Chetrit, P.; Vedel, F.
Proc. Natl. Acad. Sci. U.S.A. 90, 5934-5938, 1993
A;Title: Specific mitochondrial proteins in pollen: presence of an additional ATP syntha
A;Reference number: A48186; MUID:93317598; PMID:8327463
A;Accession: A48186
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-23 <DE1>
A;Experimental source: pollen
A;Note: sequence extracted from NCBI backbone (NCBIP:134867)

  Query Match            19.1%;  Score 27;  DB 2;  Length 23;
  Best Local Similarity   37.5%;  Pred. No. 1.2e+03;
  Matches    6;  Conservative    2;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          1 AVPIAQKSEPHSLSSE 16
              | | :| : |    ||
Db          5 AAPASQPTPPKPSGSE 20


RESULT 11
A30323
amyloid protein AL (Ig lambda chain V region) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 21-Feb-1990 #sequence_revision 21-Feb-1990 #text_change 16-Aug-1996
C;Accession: A30323
R;Gorevic, P.D.; Prelli, F.C.; Wright, J.; Pras, M.; Frangione, B.
J. Clin. Invest. 83, 836-843, 1989
A;Title: Systemic senile amyloidosis. Identification of a new prealbumin (transthyret
opathy.
A;Reference number: A30323; MUID:89155805; PMID:2646319
A;Accession: A30323
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-27 <GOR>
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin

  Query Match            19.1%;  Score 27;  DB 2;  Length 27;
  Best Local Similarity   57.1%;  Pred. No. 1.4e+03;
  Matches    4;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          8 SEPHSLS 14
              ::|||:|
Db          5 TQPHSVS 11


RESULT 12
T29626
hypothetical protein K09E3.3 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T29626
R;Johnson, D.; Gattung, S.
submitted to the EMBL Data Library, November 1995
A;Description: The sequence of C. elegans cosmid K09E3.
A;Reference number: Z20655
A;Accession: T29626
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-24 <JOH>
A;Cross-references: EMBL:U41033; PIDN:AAA82374.1; CESP:K09E3.3
C;Genetics:
A;Gene: CESP:K09E3.3

  Query Match            18.4%;  Score 26;  DB 2;  Length 24;
  Best Local Similarity   31.8%;  Pred. No. 1.8e+03;
  Matches    7;  Conservative    6;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          4 IAQKSEPHSLSSEALMRRAVSL 25
              ::|:     |:: ||: |  ||
Db          2 VSQRVCAAGLTARALLARPNSL 23


RESULT 13
A37825
fibronectin receptor alpha chain - chicken (fragment)
C;Species: Gallus gallus (chicken)
C;Date: 30-Apr-1991 #sequence_revision 30-Apr-1991 #text_change 18-Jun-1993
C;Accession: A37825
R;Hofer, U.; Syfrig, J.; Chiquet-Ehrismann, R.
J. Biol. Chem. 265, 14561-14565, 1990
A;Title: Identification and characterization of a dimeric chicken fibronectin recepto
A;Reference number: A37825; MUID:90354452; PMID:2143763
A;Accession: A37825
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-24 <HOF>

  Query Match            18.4%;  Score 26;  DB 2;  Length 24;
  Best Local Similarity   43.8%;  Pred. No. 1.8e+03;
  Matches    7;  Conservative    1;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          3 PIAQKSEPHSLSSEAL 18
              | | :  | ||   ||
Db          8 PFAFRGSPGSLFGFAL 23


RESULT 14
```

A43768
Hu-like protein HB1 - Bifidobacterium longum (fragment)
C;Species: Bifidobacterium longum
C;Date: 01-Dec-1992 #sequence_revision 01-Dec-1992 #text_change 16-Feb-1997
C;Accession: A43768
R;Goshima, N.; Kano, Y.; Imamoto, F.
Biochimie 72, 207-212, 1990
A;Title: Characterization of HU-like protein from Bifidobacterium longum.
A;Reference number: A43768; MUID:90344917; PMID:2116910
A;Accession: A43768
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-27 <GOS>
C;Keywords: DNA binding

Query Match 18.4%; Score 26; DB 2; Length 27;
Best Local Similarity 50.0%; Pred. No. 2.1e+03;
Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

```
Qy          4 IAQKSEPHSLSSEA 17
              |||||      :||
Db         11 IAQKSNLTKAQAEA 24
```

RESULT 15
A60558
cytochrome P450 HLp3 - human (fragment)
N;Contains: oxidoreductase (EC 1.-.-.-)
C;Species: Homo sapiens (man)
C;Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 17-Mar-1999
C;Accession: A60558
R;Wrighton, S.A.; Ring, B.J.; Watkins, P.B.; Vandenbranden, M.
Mol. Pharmacol. 36, 97-105, 1989
A;Title: Identification of a polymorphically expressed member of the human cytochrome P-
A;Reference number: A60558; MUID:89313723; PMID:2747634
A;Accession: A60558
A;Molecule type: protein
A;Residues: 1-29 <WRI>
C;Comment: This protein strongly resembles, but is distinct from, cytochrome P450 CYP3A5
C;Superfamily: human cytochrome P450 CYP3A5; cytochrome P450 homology
C;Keywords: electron transfer; endoplasmic reticulum; heme; monooxygenase; oxidoreductas

Query Match 18.4%; Score 26; DB 2; Length 29;
Best Local Similarity 46.7%; Pred. No. 2.2e+03;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

```
Qy         12 SLSSEALMRRAVSLV 26
              :|: |  :  |||||
Db          6 NLAVETXLLLAVSLV 20
```

Search completed: October 2, 2003, 09:43:21
Job time : 19 secs

GenCore version 5.1.6


OM protein - protein search, using sw model

Run on: October 2, 2003, 09:39:23 ; Search time 11 Seconds
(without alignments)
128.255 Million cell updates/sec

Title: US-09-939-293A-19_COPY_56_85
Perfect score: 141
Sequence: 1 AVPIAQKSEPHSLSSEALMRRAVSLVTDST 30

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 2025

Minimum DB seq length: 0
Maximum DB seq length: 30

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 34 | 24.1 | 15 | 1 | LPL_THETH | P21234 thermus the |
| 2 | 32 | 22.7 | 27 | 1 | LIPS_BOVIN | P16386 bos taurus |
| 3 | 29 | 20.6 | 25 | 1 | NCP_PIG | P80552 sus scrofa |
| 4 | 28 | 19.9 | 17 | 1 | UP35_UPEMJ | P82042 uperoleia m |
| 5 | 26 | 18.4 | 16 | 1 | FOR1_MYRGU | P81438 myrmecia gu |
| 6 | 26 | 18.4 | 29 | 1 | ATPA_BRYMA | P26965 bryopsis ma |
| 7 | 26 | 18.4 | 29 | 1 | DMD_RAT | P11530 rattus norv |
| 8 | 25.5 | 18.1 | 20 | 1 | SODF_PASPI | P81527 pasteurella |
| 9 | 25 | 17.7 | 20 | 1 | LECB_IRIHO | P36231 iris hollan |
| 10 | 25 | 17.7 | 20 | 1 | RECX_AZOVI | P37863 azotobacter |
| 11 | 25 | 17.7 | 21 | 1 | ALL7_OLEEU | P81430 olea europa |
| 12 | 25 | 17.7 | 22 | 1 | IAD1_ENTFA | P24803 enterococcu |
| 13 | 25 | 17.7 | 24 | 1 | CH60_HELVI | P26317 heliothis v |
| 14 | 25 | 17.7 | 29 | 1 | CERB_CERCA | P36191 ceratitis c |
| 15 | 25 | 17.7 | 30 | 1 | RL18_HALCU | P05970 halobacteri |
| 16 | 24.5 | 17.4 | 29 | 1 | SODC_OLEEU | P80740 olea europa |
| 17 | 24 | 17.0 | 12 | 1 | LMT1_LOCMI | P22395 locusta mig |
| 18 | 24 | 17.0 | 20 | 1 | RLC1_HALMA | P12740 haloarcula |
| 19 | 24 | 17.0 | 21 | 1 | MCT3_MOUSE | P21843 mus musculu |
| 20 | 24 | 17.0 | 24 | 1 | CH60_ACICA | P81874 acinetobact |
| 21 | 23.5 | 16.7 | 22 | 1 | LPL_CORGL | P42456 corynebacte |
| 22 | 23 | 16.3 | 17 | 1 | UP34_UPEMJ | P82041 uperoleia m |
| 23 | 23 | 16.3 | 18 | 1 | YMDB_CHLAU | Q45827 chloroflexu |
| 24 | 23 | 16.3 | 20 | 1 | H2B1_ICTPU | P81903 ictalurus p |
| 25 | 23 | 16.3 | 21 | 1 | TRYP_APIMS | P83348 apis mellif |
| 26 | 23 | 16.3 | 23 | 1 | T2B_PARTE | Q27174 paramecium |
| 27 | 23 | 16.3 | 29 | 1 | DMS5_PHYSA | P80281 phyllomedus |
| 28 | 22 | 15.6 | 15 | 1 | CKX_WHEAT | P58763 triticum ae |
| 29 | 22 | 15.6 | 15 | 1 | COXJ_THUOB | P80979 thunnus obe |
| 30 | 22 | 15.6 | 20 | 1 | COG1_CHIOP | P34153 chionoecete |
| 31 | 22 | 15.6 | 20 | 1 | DER6_DERPT | P49277 dermatophag |
| 32 | 22 | 15.6 | 20 | 1 | SCB1_CANFA | P99507 canis famil |
| 33 | 22 | 15.6 | 22 | 1 | CH60_BOVIN | P31081 bos taurus |
| 34 | 22 | 15.6 | 27 | 1 | DCUP_RHOSH | P32920 rhodobacter |
| 35 | 22 | 15.6 | 27 | 1 | IORA_METTM | P80910 methanobact |
| 36 | 22 | 15.6 | 27 | 1 | PSAF_SYNVU | P23078 synechococc |
| 37 | 22 | 15.6 | 28 | 1 | ARYA_PSEPU | P35902 pseudomonas |
| 38 | 22 | 15.6 | 28 | 1 | COXB_SOLTU | P80499 solanum tub |
| 39 | 21 | 14.9 | 12 | 1 | TM2A_METMA | P80652 methanosarc |
| 40 | 21 | 14.9 | 14 | 1 | UC15_MAIZE | P80621 zea mays (m |
| 41 | 21 | 14.9 | 15 | 1 | MAOX_CHICK | Q92060 gallus gall |
| 42 | 21 | 14.9 | 15 | 1 | PDGB_PIG | P20034 sus scrofa |
| 43 | 21 | 14.9 | 15 | 1 | UN01_PINPS | P81106 pinus pinas |
| 44 | 21 | 14.9 | 17 | 1 | H2B3_ICTPU | P81904 ictalurus p |
| 45 | 21 | 14.9 | 20 | 1 | CP35_PAPSP | P80056 papio sp. ( |

ALIGNMENTS

RESULT 1
LPL_THETH
ID LPL_THETH STANDARD; PRT; 15 AA.
AC P21234;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Leu leader peptide.
OS Thermus thermophilus.
OC Bacteria; Deinococcus-Thermus; Deinococci; Thermales; Thermaceae;
OC Thermus.
OX NCBI_TaxID=274;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HB8 / ATCC 27634;
RX MEDLINE=88121725; PubMed=3323845;
RA Croft J.E., Love D.R., Bergquist P.L.;
RT "Expression of leucine genes from an extremely thermophilic bacterium
RT in Escherichia coli.";
RL Mol. Gen. Genet. 210:490-497(1987).
CC -!- FUNCTION: Involved in control of the biosynthesis of leucine.
CC ---

CC ---
DR EMBL; X06604; CAA29823.1; -.
KW Leucine biosynthesis; Leader peptide.
SQ SEQUENCE 15 AA; 1666 MW; C2F107A386D7620B CRC64;

Query Match 24.1%; Score 34; DB 1; Length 15;
Best Local Similarity 80.0%; Pred. No. 29;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy         19 MRRAVSLVTD 28
              ||||| || |
Db          1 MRRAVILVLD 10
```

RESULT 2
LIPS_BOVIN
ID LIPS_BOVIN STANDARD; PRT; 27 AA.
AC P16386;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hormone sensitive lipase (EC 3.1.1.-) (HSL) (Fragment).
GN LIPE.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Bovinae; Bos.

```
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=88152238; PubMed=3345839;
RA   Garton A.J., Campbell D.G., Cohen P., Yeaman S.J.;
RT   "Primary structure of the site on bovine hormone-sensitive lipase
RT   phosphorylated by cyclic AMP-dependent protein kinase.";
RL   FEBS Lett. 229:68-72(1988).
RN   [2]
RP   SEQUENCE OF 8-12, AND PHOSPHORYLATION OF SER-10.
RC   TISSUE=Adipose tissue;
RX   MEDLINE=89137090; PubMed=2537200;
RA   Garton A.J., Campbell D.G., Carling D., Hardie D.G., Colbran R.J.,
RA   Yeaman S.J.;
RT   "Phosphorylation of bovine hormone-sensitive lipase by the
RT   AMP-activated protein kinase. A possible antilipolytic mechanism.";
RL   Eur. J. Biochem. 179:249-254(1989).
CC   -!- FUNCTION: IN ADIPOSE TISSUE AND HEART, IT PRIMARILY HYDROLYZES
CC       STORED TRIGLYCERIDES TO FREE FATTY ACIDS, WHILE IN STEROIDOGENIC
CC       TISSUES, IT PRINCIPALLY CONVERTS CHOLESTERYL ESTERS TO FREE
CC       CHOLESTEROL FOR STEROID HORMONE PRODUCTION.
CC   -!- ENZYME REGULATION: RAPIDLY ACTIVATED BY CAMP-DEPENDENT
CC       PHOSPHORYLATION UNDER THE INFLUENCE OF CATECHOLAMINES.
CC       DEPHOSPHORYLATION AND INACTIVATION ARE CONTROLLED BY INSULIN.
CC   -!- PATHWAY: HORMONE SENSITIVE LIPASE CATALYZES THE RATE LIMITING
CC       STEP IN TRIGLYCERIDE LIPOLYSIS.
CC   -!- SIMILARITY: BELONGS TO THE "GDXG" FAMILY OF LIPOLYTIC ENZYMES.
DR   PIR; S00347; S00347.
DR   InterPro; IPR002168; Lipolytic_enzyme.
DR   PROSITE; PS01173; LIPASE_GDXG_HIS; PARTIAL.
DR   PROSITE; PS01174; LIPASE_GDXG_SER; PARTIAL.
KW   Hydrolase; Lipid degradation; Phosphorylation.
FT   NON_TER       1      1
FT   MOD_RES       8      8       PHOSPHORYLATION (BY PKA).
FT   MOD_RES      10     10       PHOSPHORYLATION (BY AMPK).
FT   NON_TER      27     27
SQ   SEQUENCE   27 AA;  2899 MW;  7ADFA70711D71858 CRC64;

  Query Match             22.7%;  Score 32;  DB 1;  Length 27;
  Best Local Similarity   44.0%;  Pred. No. 1.2e+02;
  Matches   11;  Conservative    3;  Mismatches    9;  Indels    2;  Gaps    1;

Qy          7 KSEP--HSLSSEALMRRAVSLVTDS 29
              |:||   |:|  || :    | |||
Db          1 KTEPMRRSVSEAALTQPEGPLGTDS 25


RESULT 3
NCP_PIG
ID   NCP_PIG        STANDARD;      PRT;    25 AA.
AC   P80552;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   20 kDa neutrophil cationic protein (NCP) (Fragment).
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=96242065; PubMed=8645990;
RA   Fornhem C., Peterson C.G.B., Alving K.;
RT   "Isolation and characterization of porcine cationic eosinophil
RT   granule proteins.";
RL   Int. Arch. Allergy Immunol. 110:132-142(1996).
FT   NON_TER      25     25
SQ   SEQUENCE   25 AA;  2629 MW;  5275BFFF8D81F3AD CRC64;

  Query Match             20.6%;  Score 29;  DB 1;  Length 25;
  Best Local Similarity   40.0%;  Pred. No. 3.1e+02;
  Matches    6;  Conservative    4;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 VPIAQKSEPHSLSSE 16
              :||  : |  :|:||
Db          6 IPIVSRREWGALASE 20


RESULT 4
UP35_UPEMJ
ID   UP35_UPEMJ     STANDARD;      PRT;    17 AA.
AC   P82042;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Uperin 3.5.
OS   Uperoleia mjobergii (Australian toadlet).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Myobatrachidae;
OC   Myobatrachinae; Uperoleia.
OX   NCBI_TaxID=104954;
RN   [1]
RP   SEQUENCE, AND MASS SPECTROMETRY.
RC   TISSUE=Skin secretion;
RA   Bradford A.M., Bowie J.H., Tyler M.J., Wallace J.C.;
RT   "New antibiotic uperin peptides from the dorsal glands of the
RT   australian toadlet Uperoleia mjobergii.";
RL   Aust. J. Chem. 49:1325-1331(1996).
CC   -!- FUNCTION: SHOWS ANTIBACTERIAL ACTIVITY AGAINST B.CEREUS, L.LACTIS,
CC       L.INNOCUA, M.LUTEUS, S.AUREUS, P.MULTOCI, S.EPIDERMIS AND
CC       S.UBERIS.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Expressed by the skin dorsal glands.
CC   -!- MASS SPECTROMETRY: MW=1779; METHOD=FAB.
KW   Amphibian defense peptide; Antibiotic; Amidation.
FT   MOD_RES      17     17       AMIDATION.
SQ   SEQUENCE   17 AA;  1781 MW;  6F61E483436AD67B CRC64;

  Query Match             19.9%;  Score 28;  DB 1;  Length 17;
  Best Local Similarity   55.6%;  Pred. No. 2.8e+02;
  Matches    5;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         18 LMRRAVSLV 26
              |:|:|||::
Db          5 LIRKAVSVI 13


RESULT 5
FOR1_MYRGU
ID   FOR1_MYRGU     STANDARD;      PRT;    16 AA.
AC   P81438;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   Formaecin 1.
OS   Myrmecia gulosa (Red bulldog ant).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Formicidae;
OC   Myrmecinae; Myrmecia.
OX   NCBI_TaxID=36170;
RN   [1]
RP   SEQUENCE, AND CARBOHYDRATE-LINKAGE SITE THR-11.
RC   TISSUE=Hemolymph;
RX   MEDLINE=98165787; PubMed=9497332;
RA   Mackintosh J.A., Veal D.A., Beattie A.J., Gooley A.A.;
RT   "Isolation from an ant Myrmecia gulosa of two inducible
RT   O-glycosylated proline-rich antibacterial peptides.";
RL   J. Biol. Chem. 273:6139-6143(1998).
CC   -!- FUNCTION: ANTIBACTERIAL PEPTIDE. HAS ACTIVITY AGAINST E.COLI
CC       BUT NONE AGAINST OTHER GRAM-NEGATIVE BACTERIA AND GRAM-POSITIVE
CC       BACTERIA.
CC   -!- INDUCTION: By bacterial infection.
CC   -!- PTM: O-LINKED GLYCAN CONSISTS OF A GAL-GALNAC DISACCHARIDE, O-
CC       GLYCOSYLATION IS ESSENTIAL FOR FULL BIOLOGICAL ACTIVITY.
CC   -!- SIMILARITY: TO DROSOPHILA DROSOCIN.
KW   Antibiotic; Glycoprotein; Insect immunity; Hemolymph.
```

```
FT   CARBOHYD     11     11       O-LINKED (GALNAC. . .).
SQ   SEQUENCE   16 AA;  1794 MW;  80CEA3AABBC2E0AE CRC64;

  Query Match             18.4%;  Score 26;  DB 1;  Length 16;
  Best Local Similarity   44.4%;  Pred. No. 5.2e+02;
  Matches    4;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          3 PIAQKSEPH 11
              |:  |  ||
Db          5 PVNNKPTPH 13


RESULT 6
ATPA_BRYMA
ID   ATPA_BRYMA     STANDARD;      PRT;    29 AA.
AC   P26965;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   ATP synthase alpha chain (EC 3.6.3.14) (Fragment).
GN   ATPA.
OS   Bryopsis maxima (Green alga).
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Chlorophyta; Ulvophyceae; Caulerpales;
OC   Bryopsidaceae; Bryopsis.
OX   NCBI_TaxID=3129;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91355942; PubMed=1884001;
RA   Kono M., Satoh H., Okabe Y., Abe Y., Nakayama K., Okada M.;
RT   "Nucleotide sequence of the large subunit of
RT   ribulose-1,5-bisphosphate carboxylase/oxygenase from the green alga
RT   Bryopsis maxima.";
RL   Plant Mol. Biol. 17:505-508(1991).
CC   -!- FUNCTION: PRODUCES ATP FROM ADP IN THE PRESENCE OF A PROTON
CC       GRADIENT ACROSS THE MEMBRANE. THE ALPHA CHAIN IS A REGULATORY
CC       SUBUNIT.
CC   -!- CATALYTIC ACTIVITY: ATP + H(2)O + H(+)(In) = ADP + phosphate +
CC       H(+)(Out).
CC   -!- SUBUNIT: F-TYPE ATPASES HAVE 2 COMPONENTS, CF(1) - THE CATALYTIC
CC       CORE - AND CF(0) - THE MEMBRANE PROTON CHANNEL. CF(1) HAS FIVE
CC       SUBUNITS: ALPHA(3), BETA(3), GAMMA(1), DELTA(1), EPSILON(1). CF(0)
CC       HAS THREE MAIN SUBUNITS: A, B AND C.
CC   -!- SUBCELLULAR LOCATION: Chloroplast thylakoid membrane.
CC   -!- SIMILARITY: Belongs to the ATPase alpha/beta chains family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X55877; CAA39362.1; -.
DR   PIR; S17432; S17432.
DR   InterPro; IPR000194; ATPase_a/bcentre.
DR   PROSITE; PS00152; ATPASE_ALPHA_BETA; PARTIAL.
KW   ATP synthesis; Chloroplast; Thylakoid; Membrane; CF(1);
KW   ATP-binding; Hydrolase; Hydrogen ion transport.
FT   NON_TER       1      1
SQ   SEQUENCE   29 AA;  3308 MW;  A25A0BAD077F338B CRC64;

  Query Match             18.4%;  Score 26;  DB 1;  Length 29;
  Best Local Similarity   30.0%;  Pred. No. 1.1e+03;
  Matches    6;  Conservative    5;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          4 IAQKSEPHSLSSEALMRRAV 23
              |   :   |  :|||:: |:
Db          2 IIMSTNTFSEEAEALLKEAL 21


RESULT 7
```

```
DMD_RAT
ID   DMD_RAT        STANDARD;      PRT;    29 AA.
AC   P11530;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Dystrophin (Fragment).
GN   DMD.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88122671; PubMed=3340214;
RA   Nudel U., Robzyk K., Yaffe D.;
RT   "Expression of the putative Duchenne muscular dystrophy gene in
RT   differentiated myogenic cell cultures and in the brain.";
RL   Nature 331:635-638(1988).
CC   -!- FUNCTION: May play a role in anchoring the cytoskeleton to the
CC       plasma membrane.
CC   -!- SUBUNIT: Interacts with the syntrophins SNTA1, SNTB1, SNTB2, SNTG1
CC       and SNTG2 (By similarity).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X07000; CAA30057.1; -.
DR   PIR; S01614; S01614.
DR   InterPro; IPR001589; Actbind_actnin.
DR   InterPro; IPR001202; WW_Rsp5_WWP.
DR   PROSITE; PS00019; ACTININ_1; PARTIAL.
DR   PROSITE; PS00020; ACTININ_2; PARTIAL.
DR   PROSITE; PS01159; WW_DOMAIN_1; PARTIAL.
DR   PROSITE; PS50020; WW_DOMAIN_2; PARTIAL.
KW   Structural protein; Actin-binding; Calcium-binding; Cytoskeleton;
KW   Repeat.
FT   NON_TER       1      1
FT   NON_TER      29     29
SQ   SEQUENCE   29 AA;  3289 MW;  8ECFB28A1A7ACAF0 CRC64;

  Query Match             18.4%;  Score 26;  DB 1;  Length 29;
  Best Local Similarity   31.6%;  Pred. No. 1.1e+03;
  Matches    6;  Conservative    7;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          6 QKSEPHSLSSEALMRRAVS 24
              :| :  | |::||: : |:
Db         11 RKLQDASRSAQALVEQMVN 29


RESULT 8
SODF_PASPI
ID   SODF_PASPI     STANDARD;      PRT;    20 AA.
AC   P81527;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Superoxide dismutase [Fe] (EC 1.15.1.1) (Fragment).
GN   SODB.
OS   Pasteurella piscicida (Photobacterium damsela (subsp. piscicida)).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC   Vibrionaceae; Photobacterium.
OX   NCBI_TaxID=38294;
RN   [1]
RP   SEQUENCE.
RC   STRAIN=MT1415;
RX   MEDLINE=99173752; PubMed=10075430;
RA   Barnes A.C., Balebona M.C., Horne M.T., Ellis A.E.;
```

```
RT   "Superoxide dismutase and catalase in Photobacterium damselae subsp.
RT   piscicida and their roles in resistance to reactive oxygen species.";
RL   Microbiology 145:483-494(1999).
CC   -!- FUNCTION: Destroys radicals which are normally produced within the
CC       cells and which are toxic to biological systems.
CC   -!- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
CC   -!- COFACTOR: Iron.
CC   -!- SUBUNIT: Homodimer (By similarity).
CC   -!- SUBCELLULAR LOCATION: Periplasmic.
CC   -!- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE
CC       FAMILY.
DR   InterPro; IPR001189; SODismutase.
DR   Pfam; PF00081; sodfe; 1.
KW   Oxidoreductase; Iron; Periplasmic.
FT   NON_TER      20     20
SQ   SEQUENCE   20 AA;  2153 MW;  A8D31FDAE8553B6D CRC64;

  Query Match             18.1%;  Score 25.5;  DB 1;  Length 20;
  Best Local Similarity   50.0%;  Pred. No. 8.1e+02;
  Matches    6;  Conservative    3;  Mismatches    2;  Indels    1;  Gaps    1;

Qy          1 AVPIAQKS-EPH 11
              |:| |: : |||
Db          6 ALPYAKDALEPH 17


RESULT 9
LECB_IRIHO
ID   LECB_IRIHO     STANDARD;      PRT;     20 AA.
AC   P36231;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   N-acetyl-D-galactosamine-binding lectin subunit B (A-disaccharide-
DE   binding lectin subunit B) (Fragment).
OS   Iris hollandica (Dutch iris).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Iridaceae;
OC   Iris.
OX   NCBI_TaxID=35876;
RN   [1]
RP   SEQUENCE.
RC   STRAIN=cv. Golden Harvest, and cv. Prof. Blaauw; TISSUE=Bulb;
RX   MEDLINE=94171801; PubMed=8125993;
RA   Mo H., van Damme E.J.M., Peumans W.J., Goldstein I.J.;
RT   "Isolation and characterization of an
RT   N-acetyl-D-galactosamine-binding lectin from Dutch Iris bulbs which
RT   recognizes the blood group A disaccharide (GalNAc alpha 1-3Gal).";
RL   J. Biol. Chem. 269:7666-7673(1994).
CC   -!- FUNCTION: GAL / GALNAC-SPECIFIC LECTIN. AGGLUTINATES BOTH NATIVE
CC       AND TRYPSIN-TREATED RABBIT ERYTHROCYTES BUT NOT HUMAN ERYTHROCYTES
CC       IRRESPECTIVE OF BLOOD GROUP TYPE.
CC   -!- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A AND B CHAINS.
KW   Lectin.
FT   NON_TER      20     20
SQ   SEQUENCE   20 AA;  2246 MW;  DE08DA025FD17D56 CRC64;

  Query Match             17.7%;  Score 25;  DB 1;  Length 20;
  Best Local Similarity   62.5%;  Pred. No. 9.7e+02;
  Matches    5;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         13 LSSEALMR 20
              || | :||
Db          7 LSEETMMR 14


RESULT 10
RECX_AZOVI
ID   RECX_AZOVI     STANDARD;      PRT;     20 AA.
AC   P37863;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Regulatory protein recX (Fragment).
GN   RECX.
OS   Azotobacter vinelandii.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC   Pseudomonadaceae; Azotobacter.
OX   NCBI_TaxID=354;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92225347; PubMed=1563632;
RA   Venkatesh T.V., Das H.K.;
RT   "The Azotobacter vinelandii recA gene: sequence analysis and
RT   regulation of expression.";
RL   Gene 113:47-53(1992).
RN   [2]
RP   IDENTIFICATION.
RX   MEDLINE=94218258; PubMed=8165147;
RA   de Mot R., Schoofs G., Vanderleyden J.;
RT   "A putative regulatory gene downstream of recA is conserved in gram-
RT   negative and Gram-positive bacteria.";
RL   Nucleic Acids Res. 22:1313-1314(1994).
CC   -!- FUNCTION: Modulates recA activity (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC   -!- SIMILARITY: BELONGS TO THE RECX FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; S96898; -; NOT_ANNOTATED_CDS.
DR   HAMAP; MF_01114; -; 1.
FT   NON_TER      20     20
SQ   SEQUENCE   20 AA;  2111 MW;  C809F8BCCED6CB56 CRC64;

  Query Match             17.7%;  Score 25;  DB 1;  Length 20;
  Best Local Similarity   60.0%;  Pred. No. 9.7e+02;
  Matches    6;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy         13 LSSEALMRRA 22
              | | | :|||
Db          4 LDSPAAVRRA 13


RESULT 11
ALL7_OLEEU
ID   ALL7_OLEEU     STANDARD;      PRT;     21 AA.
AC   P81430;
DT   30-MAY-2000 (Rel. 39, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Pollen allergen Ole e 7 (Ole e VII) (Fragment).
OS   Olea europaea (Common olive).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; lamiids; Lamiales; Oleaceae; Olea.
OX   NCBI_TaxID=4146;
RN   [1]
RP   SEQUENCE (VARIANTS A AND B), AND MASS SPECTROMETRY.
RC   TISSUE=Pollen;
RX   MEDLINE=99449676; PubMed=10518824;
RA   Tejera M.L., Villalba M., Batanero E., Rodriguez R.;
RT   "Identification, isolation, and characterization of Ole e 7, a new
RT   allergen of olive tree pollen.";
RL   J. Allergy Clin. Immunol. 104:797-802(1999).
CC   -!- POLYMORPHISM: Many isoforms of the allergen exist due to
CC       polymorphism. They can be classified as isoforms of type A (shown
CC       here) and isoforms of type B. A microheterogeneity is detected at
CC       positions 4 and 11 of isoforms of type A and at positions 4, 5, 10
CC       and 11 of isoforms of type B.
CC   -!- MISCELLANEOUS: Allergen from olive pollen. Important in
```

```
CC       mediterranean countries and California. Its prevalence is related
CC       to the geographic area.
KW   Allergen; Polymorphism.
FT   VARIANT       5      5       S -> G (IN TYPE B).
FT   VARIANT      10     10       L -> K (IN TYPE B).
FT   VARIANT      18     18       I -> K (IN TYPE B).
FT   NON_TER      21     21
SQ   SEQUENCE   21 AA;  2199 MW;  F0E9B99FEB079400 CRC64;

  Query Match             17.7%;  Score 25;  DB 1;  Length 21;
  Best Local Similarity   31.6%;  Pred. No. 1e+03;
  Matches    6;  Conservative    3;  Mismatches   10;  Indels    0;  Gaps    0;

Qy         10 PHSLSSEALMRRAVSLVTD 28
              |    :  ||:   || : |
Db          2 PSQSTVTALLTSCVSYIDD 20


RESULT 12
IAD1_ENTFA
ID   IAD1_ENTFA     STANDARD;      PRT;    22 AA.
AC   P24803;
DT   01-MAR-1992 (Rel. 21, Created)
DT   01-MAR-1992 (Rel. 21, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Sex pheromone inhibitor determinant precursor (iAD1).
GN   IAD OR EFA0005.1.
OS   Enterococcus faecalis (Streptococcus faecalis).
OG   Plasmid pTEF1, and Plasmid pAD1.
OC   Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
OX   NCBI_TaxID=1351;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   PLASMID=pAD1;
RX   MEDLINE=91261999; PubMed=2128961;
RA   Clewell D.B., Pontius L.T., An F.Y., Ike Y., Suzuki A., Nakayama J.;
RT   "Nucleotide sequence of the sex pheromone inhibitor (iAD1)
RT   determinant of Enterococcus faecalis conjugative plasmid pAD1.";
RL   Plasmid 24:156-161(1990).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=V583 / ATCC 700802; PLASMID=pTEF1;
RX   MEDLINE=22550857; PubMed=12663927;
RA   Paulsen I.T., Banerjei L., Myers G.S.A., Nelson K.E., Seshadri R.,
RA   Read T.D., Fouts D.E., Eisen J.A., Gill S.R., Heidelberg J.F.,
RA   Tettelin H., Dodson R.J., Umayam L., Brinkac L., Beanan M.,
RA   Daugherty S., DeBoy R.T., Durkin S., Kolonay J., Madupu R., Nelson W.,
RA   Vamathevan J., Tran B., Upton J., Hansen T., Shetty J., Khouri H.,
RA   Utterback T., Radune D., Ketchum K.A., Dougherty B.A., Fraser C.M.;
RT   "Role of mobile DNA in the evolution of vancomycin-resistant
RT   Enterococcus faecalis.";
RL   Science 299:2071-2074(2003).
CC   -!- FUNCTION: ACTS AS A COMPETITIVE INHIBITOR OF THE CAD1 PHEROMONE.
CC   -!- SUBCELLULAR LOCATION: Secreted (Probable).
CC   -!- MISCELLANEOUS: IAD1 APPEARS TO BE A COMPONENT OF ITS OWN SIGNAL
CC       SEQUENCE.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; M62888; AAA98039.1; -.
DR   EMBL; AE016833; AAO83007.1; -.
DR   PIR; A37391; A37391.
DR   TIGR; EFA0005.1; -.
KW   Plasmid.
FT   PROPEP        1     14       POTENTIAL.
FT   CHAIN        15     22       SEX PHEROMONE INHIBITOR DETERMINANT.
SQ   SEQUENCE   22 AA;  2459 MW;  D0EAEBDF1BCD9D08 CRC64;

  Query Match             17.7%;  Score 25;  DB 1;  Length 22;
  Best Local Similarity   30.8%;  Pred. No. 1.1e+03;
  Matches    4;  Conservative    5;  Mismatches    4;  Indels    0;  Gaps    0;

Qy         15 SEALMRRAVSLVT 27
              |:  |:: : |:|
Db          2 SKRAMKKIIPLIT 14


RESULT 13
CH60_HELVI
ID   CH60_HELVI     STANDARD;      PRT;    24 AA.
AC   P26317;
DT   01-MAY-1992 (Rel. 22, Created)
DT   01-MAY-1992 (Rel. 22, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   60 kDa chaperonin, mitochondrial (P60) (Fragment).
OS   Heliothis virescens (Noctuid moth) (Owlet moth).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Noctuoidea;
OC   Noctuidae; Heliothinae; Heliothis.
OX   NCBI_TaxID=7102;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Testis;
RX   MEDLINE=90339485; PubMed=1974308;
RA   Miller S.G., Leclerc R.F., Erdos G.W.;
RT   "Identification and characterization of a testis-specific isoform of
RT   a chaperonin in a moth, Heliothis virescens.";
RL   J. Mol. Biol. 214:407-422(1990).
CC   -!- FUNCTION: IMPLICATED IN MITOCHONDRIAL PROTEIN IMPORT AND
CC       MACROMOLECULAR ASSEMBLY. MAY FACILITATE THE CORRECT FOLDING OF
CC       IMPORTED PROTEINS. MAY ALSO PREVENT MISFOLDING AND PROMOTE THE
CC       REFOLDING AND PROPER ASSEMBLY OF UNFOLDED POLYPEPTIDES GENERATED
CC       UNDER STRESS CONDITIONS IN THE MITOCHONDRIAL MATRIX (BY
CC       SIMILARITY).
CC   -!- SUBUNIT: FORMS A SINGLE SEVEN-MEMBER RING COMPLEX, IN TIGHT
CC       ASSOCIATION WITH THE P63 PROTEIN.
CC   -!- SUBCELLULAR LOCATION: Mitochondrial.
CC   -!- TISSUE SPECIFICITY: Testis.
CC   -!- DEVELOPMENTAL STAGE: FROM THE SECOND HALF OF THE LARVAL FINAL-
CC       INSTAR, THROUGH THE FIRST TWO DAYS OF PUPAL DEVELOPMENT.
CC   -!- MISCELLANEOUS: SHOWS ATPASE ACTIVITY.
CC   -!- SIMILARITY: Belongs to the chaperonin (HSP60) family.
DR   InterPro; IPR001844; Chaprnin_Cpn60.
DR   PROSITE; PS00296; CHAPERONINS_CPN60; PARTIAL.
KW   Chaperone; ATP-binding; Testis; Mitochondrion.
FT   NON_TER      24     24
SQ   SEQUENCE   24 AA;  2531 MW;  2B34508F8CA981CF CRC64;

  Query Match             17.7%;  Score 25;  DB 1;  Length 24;
  Best Local Similarity   38.5%;  Pred. No. 1.2e+03;
  Matches    5;  Conservative    4;  Mismatches    4;  Indels    0;  Gaps    0;

Qy         17 ALMRRAVSLVTDS 29
              ||| : | :: |:
Db         12 ALMLQGVDVLADA 24


RESULT 14
CERB_CERCA
ID   CERB_CERCA     STANDARD;      PRT;    29 AA.
AC   P36191;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-FEB-1996 (Rel. 33, Last annotation update)
DE   Ceratotoxin B.
GN   CTXB.
OS   Ceratitis capitata (Mediterranean fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Tephritoidea; Tephritidae; Ceratitis.
```

```
OX   NCBI_TaxID=7213;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Female accessory gland;
RX   MEDLINE=93357786; PubMed=8353519;
RA   Marchini D., Giordano P.C., Amons R., Bernini L.F., Dallai R.;
RT   "Purification and primary structure of ceratotoxin A and B, two
RT   antibacterial peptides from the female reproductive accessory glands
RT   of the medfly Ceratitis capitata (Insecta:Diptera).";
RL   Insect Biochem. Mol. Biol. 23:591-598(1993).
CC   -!- FUNCTION: FEMALE-SPECIFIC PEPTIDES WITH POTENT ACTIVITY AGAINST
CC       GRAM-POSITIVE AND GRAM-NEGATIVE BACTERIA. THEY HAVE AS WELL
CC       HEMOLYTIC ACTIVITY. THESE PROTEINS ARE STABLE EVEN AT 100 DEGREES
CC       CELSIUS.
CC   -!- SUBUNIT: HOMOPOLYMER OF FOUR TO SIX SUBUNITS.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: STRUCTURALLY RELATED TO CECROPINS, DEFENSINS AND
CC       APIADECINS.
DR   PIR; B61613; B61613.
KW   Insect immunity; Hemolysis; Antibiotic.
SQ   SEQUENCE   29 AA;  2861 MW;  EE57F4EECB2DA6B0 CRC64;

  Query Match             17.7%;  Score 25;  DB 1;  Length 29;
  Best Local Similarity   57.1%;  Pred. No. 1.5e+03;
  Matches    4;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AVPIAQK 7
              |:|:|:|
Db          9 ALPVAKK 15


RESULT 15
RL18_HALCU
ID   RL18_HALCU     STANDARD;      PRT;    30 AA.
AC   P05970;
DT   01-NOV-1988 (Rel. 09, Created)
DT   01-NOV-1988 (Rel. 09, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   50S ribosomal protein L18P (HCUL18) (HL13) (Fragment).
GN   RPL18P.
OS   Halobacterium cutirubrum.
OC   Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC   Halobacteriaceae; Halobacterium.
OX   NCBI_TaxID=2242;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=79045279; PubMed=152199;
RA   Smith N., Matheson A.T., Yaguchi M., Willick G., Nazar R.N.;
RT   "The 5-S RNA-protein complex from an extreme halophile,
RT   Halobacterium cutirubrum. Purification and characterization.";
RL   Eur. J. Biochem. 89:501-509(1978).
CC   -!- SIMILARITY: BELONGS TO THE L18P FAMILY OF RIBOSOMAL PROTEINS.
DR   PIR; S07217; S07217.
KW   Ribosomal protein.
FT   NON_TER      30     30
SQ   SEQUENCE   30 AA;  3624 MW;  3A50079B1569CB74 CRC64;

  Query Match             17.7%;  Score 25;  DB 1;  Length 30;
  Best Local Similarity   27.3%;  Pred. No. 1.5e+03;
  Matches    6;  Conservative    7;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          2 VPIAQKSEPHSLSSEALMRRAV 23
              ||: :: |  :   : |: :||
Db          8 VPMRRRREVRTDYHQRLLLKAV 29


Search completed: October  2, 2003, 09:42:15
Job time : 12 secs
```

GenCore version 5.1.6


OM protein - protein search, using sw model

Run on: October 2, 2003, 09:40:13 ; Search time 31 Seconds
(without alignments)
249.728 Million cell updates/sec

Title: US-09-939-293A-19_COPY_56_85
Perfect score: 141
Sequence: 1 AVPIAQKSEPHSLSSEALMRRAVSLVTDST 30

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 16442

Minimum DB seq length: 0
Maximum DB seq length: 30

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 35 | 24.8 | 20 | 2 | P96173 | P96173 vibrio sp. |
| 2 | 35 | 24.8 | 23 | 6 | Q29399 | Q29399 oryctolagus |
| 3 | 34 | 24.1 | 27 | 13 | O57546 | O57546 lampetra pl |
| 4 | 34 | 24.1 | 30 | 2 | Q9JMV3 | Q9jmv3 escherichia |
| 5 | 33 | 23.4 | 30 | 5 | Q8IA94 | Q8ia94 caenorhabdi |
| 6 | 32 | 22.7 | 30 | 16 | Q8CLA1 | Q8cla1 yersinia pe |
| 7 | 31 | 22.0 | 20 | 2 | Q9R4W6 | Q9r4w6 helicobacte |
| 8 | 30 | 21.3 | 27 | 13 | O57547 | O57547 lampetra pl |
| 9 | 30 | 21.3 | 30 | 8 | Q9T2P9 | Q9t2p9 narcissus p |
| 10 | 29 | 20.6 | 18 | 4 | Q8WXC8 | Q8wxc8 homo sapien |
| 11 | 29 | 20.6 | 24 | 4 | Q8IUA3 | Q8iua3 homo sapien |
| 12 | 29 | 20.6 | 27 | 8 | Q9GB51 | Q9gb51 zosterops n |
| 13 | 29 | 20.6 | 27 | 8 | Q9GB48 | Q9gb48 zosterops m |
| 14 | 29 | 20.6 | 27 | 8 | Q9GB55 | Q9gb55 zosterops s |
| 15 | 29 | 20.6 | 27 | 8 | Q9GB45 | Q9gb45 zosterops p |
| 16 | 29 | 20.6 | 27 | 8 | Q9G1I8 | Q9g1i8 zosterops c |
| 17 | 29 | 20.6 | 27 | 8 | Q9G1I7 | Q9g1i7 zosterops j |
| 18 | 29 | 20.6 | 27 | 8 | Q9GB60 | Q9gb60 zosterops r |
| 19 | 29 | 20.6 | 27 | 8 | Q9GB41 | Q9gb41 zosterops k |
| 20 | 29 | 20.6 | 27 | 8 | Q9GB43 | Q9gb43 zosterops s |
| 21 | 29 | 20.6 | 27 | 8 | Q9GB58 | Q9gb58 zosterops p |
| 22 | 29 | 20.6 | 29 | 4 | Q9UCR6 | Q9ucr6 homo sapien |
| 23 | 28 | 19.9 | 10 | 11 | Q8CJE0 | Q8cje0 rattus norv |
| 24 | 28 | 19.9 | 12 | 8 | Q8HB27 | Q8hb27 picea glauc |
| 25 | 28 | 19.9 | 12 | 8 | Q8HB26 | Q8hb26 picea maria |
| 26 | 28 | 19.9 | 12 | 8 | Q8HB25 | Q8hb25 picea ruben |
| 27 | 28 | 19.9 | 13 | 2 | Q9AIR1 | Q9air1 pseudomonas |
| 28 | 28 | 19.9 | 16 | 2 | Q8L1Y7 | Q8l1y7 plectonema |
| 29 | 28 | 19.9 | 16 | 2 | Q8L1Y8 | Q8l1y8 oscillatori |
| 30 | 28 | 19.9 | 24 | 11 | Q9QUV8 | Q9quv8 mus sp. can |
| 31 | 27 | 19.1 | 20 | 2 | Q9X629 | Q9x629 unidentifie |
| 32 | 27 | 19.1 | 20 | 2 | Q9X632 | Q9x632 pseudomonas |
| 33 | 27 | 19.1 | 20 | 2 | Q9X634 | Q9x634 serratia ma |
| 34 | 27 | 19.1 | 20 | 2 | Q9WVU7 | Q9wvu7 escherichia |
| 35 | 27 | 19.1 | 20 | 2 | Q9X630 | Q9x630 leclercia a |
| 36 | 27 | 19.1 | 21 | 15 | Q87577 | Q87577 chimpanzee |
| 37 | 27 | 19.1 | 21 | 15 | Q87581 | Q87581 chimpanzee |
| 38 | 27 | 19.1 | 21 | 15 | Q87579 | Q87579 chimpanzee |
| 39 | 27 | 19.1 | 22 | 4 | Q8TDJ4 | Q8tdj4 homo sapien |
| 40 | 27 | 19.1 | 23 | 8 | Q9T2S6 | Q9t2s6 nicotiana s |
| 41 | 27 | 19.1 | 23 | 10 | Q9S8D9 | Q9s8d9 zea mays (m |
| 42 | 27 | 19.1 | 25 | 2 | Q9X639 | Q9x639 unidentifie |
| 43 | 27 | 19.1 | 25 | 2 | Q9X642 | Q9x642 leclercia a |
| 44 | 27 | 19.1 | 25 | 2 | Q9X641 | Q9x641 citrobacter |
| 45 | 27 | 19.1 | 25 | 2 | Q9WVZ7 | Q9wvz7 escherichia |

ALIGNMENTS

RESULT 1
P96173
ID P96173 PRELIMINARY; PRT; 20 AA.
AC P96173;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Leader peptide.
OS Vibrio sp. (strain 2693).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=79682;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2693;
RX MEDLINE=98274751; PubMed=9611817;
RA Xu Y., Zhang Y., Liang Z.Y., Van de Casteele M., Legrain C.,
RA Glansdorff N.;
RT "Aspartate carbamoyltransferase from a psychrophilic deep-sea
RT bacterium, Vibrio strain 2693: properties of the enzyme, genetic
RT organization and synthesis in Escherichia coli.";
RL Microbiology 144:1435-1441(1998).
DR EMBL; Y09786; CAA70922.1; -.
SQ SEQUENCE 20 AA; 2241 MW; 35C31F588FBB5D63 CRC64;

Query Match 24.8%; Score 35; DB 2; Length 20;
Best Local Similarity 53.3%; Pred. No. 1.9e+02;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

```
Qy          6 QKSEPHSLSSEALMR 20
              |:: | |||  |:|
Db          2 QRAAPSSLSSFKLVR 16
```

RESULT 2
Q29399
ID Q29399 PRELIMINARY; PRT; 23 AA.
AC Q29399;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)

```
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   Cystic fibrosis transmembrane conductance regulator (Fragment).
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93152187; PubMed=7678968;
RA   McGrath S.A., Basu A., Zeitlin P.L.;
RT   "Cystic fibrosis gene and protein expression during fetal lung
RT   development.";
RL   Am. J. Respir. Cell Mol. Biol. 8:201-208(1993).
DR   EMBL; S54552; AAB25301.1; -.
KW   Transmembrane.
FT   NON_TER       1      1
SQ   SEQUENCE   23 AA;  2575 MW;  93C44F5789AF5F75 CRC64;

  Query Match             24.8%;  Score 35;  DB 6;  Length 23;
  Best Local Similarity   53.3%;  Pred. No. 2.3e+02;
  Matches    8;  Conservative    3;  Mismatches    4;  Indels    0;  Gaps    0;

Qy         15 SEALMRRAVSLVTDS 29
              |:| : | :||| ||
Db          5 SDASIERRLSLVPDS 19


RESULT 3
O57546
ID   O57546      PRELIMINARY;      PRT;    27 AA.
AC   O57546;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Homeobox protein LpHox4-7A (Fragment).
OS   Lampetra planeri (Brook lamprey).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
OC   Petromyzontiformes; Petromyzontidae; Lampetra.
OX   NCBI_TaxID=7750;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98358009; PubMed=9694633;
RA   Sharman A.C., Holland P.W.;
RT   "Estimation of Hox gene cluster number in lampreys.";
RL   Int. J. Dev. Biol. 42:617-620(1998).
DR   EMBL; AF044802; AAC03006.1; -.
FT   NON_TER       1      1
FT   NON_TER      27     27
SQ   SEQUENCE   27 AA;  2963 MW;  65103946106203C7 CRC64;

  Query Match             24.1%;  Score 34;  DB 13;  Length 27;
  Best Local Similarity   35.0%;  Pred. No. 3.9e+02;
  Matches    7;  Conservative    3;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          3 PIAQKSEPHSLSSEALMRRA 22
              |:  :: ||     ||  ||
Db          4 PLPDEATPHRGGERALPERA 23


RESULT 4
Q9JMV3
ID   Q9JMV3      PRELIMINARY;      PRT;    30 AA.
AC   Q9JMV3;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Luciferase alpha-subunit (Fragment).
GN   LUXA.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=HB101;
RA   Lotz W., Bauer T.;
RT   "luxAB/kan-cassette for site-directed insertion mutagenesis and
RT   bacterial transcription studies.";
RL   Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=HB101;
RA   Olsson O., Koncz C., Szalay A.;
RT   "The use of luxA gene ofthe bacterial luciferase operon as a reporter
RT   gene.";
RL   Mol. Gen. Genet. 215:1-9(1998).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=HB101;
RX   MEDLINE=92114868; PubMed=1685011;
RA   Escher A., O'Kane D.J., Szalay A.;
RT   "The beta subunit polypeptide of Vibrio harveyi luciferase determines
RT   light emission at 42 degrees C.";
RL   Mol. Gen. Genet. 230:385-393(1991).
DR   EMBL; AJ249443; CAB96206.1; -.
DR   HSSP; P07740; 1LUC.
DR   InterPro; IPR002103; Bac_luciferase.
DR   Pfam; PF00296; bac_luciferase; 1.
FT   NON_TER      30     30
SQ   SEQUENCE   30 AA;  3454 MW;  2FC87235BDBE72FD CRC64;

  Query Match             24.1%;  Score 34;  DB 2;  Length 30;
  Best Local Similarity   41.2%;  Pred. No. 4.4e+02;
  Matches    7;  Conservative    4;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          9 EPHSLSSEALMRRAVSL 25
              :|  ||   :|:| |:|
Db         11 QPPELSQTEVMKRLVNL 27


RESULT 5
Q8IA94
ID   Q8IA94      PRELIMINARY;      PRT;    30 AA.
AC   Q8IA94;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical protein Y82E9BR.3b.
GN   Y82E9BR.3.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RX   MEDLINE=99069613; PubMed=9851916;
RA   Waterston R.;
RT   "Genome sequence of the nematode C. elegans: a platform for
RT   investigating biology. The C. elegans Sequencing Consortium.";
RL   Science 282:2012-2018(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Madsen C., Kalicki J., Yoakum M.;
RT   "The sequence of C. elegans cosmid Y82E9BR.";
RL   Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Waterston R.;
RL   Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC090999; AAO12396.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   30 AA;  3335 MW;  11CC9BD2869F9EA6 CRC64;
```

```
  Query Match             23.4%;  Score 33;  DB 5;  Length 30;
  Best Local Similarity   85.7%;  Pred. No. 6.3e+02;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          9 EPHSLSS 15
              ||||||:
Db         16 EPHSLSA 22


RESULT 6
Q8CLA1
ID   Q8CLA1      PRELIMINARY;      PRT;    30 AA.
AC   Q8CLA1;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical.
GN   Y1754.
OS   Yersinia pestis.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Yersinia.
OX   NCBI_TaxID=632;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=KIM5 / Biovar Mediaevalis;
RX   MEDLINE=22137863; PubMed=12142430;
RA   Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
RA   Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
RA   Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
RA   Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
RA   Perry R.D.;
RT   "Genome sequence of Yersinia pestis KIM.";
RL   J. Bacteriol. 184:4601-4611(2002).
DR   EMBL; AE013778; AAM85323.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   30 AA;  3461 MW;  2DB0CEA2207EBC7C CRC64;

  Query Match             22.7%;  Score 32;  DB 16;  Length 30;
  Best Local Similarity   77.8%;  Pred. No. 9e+02;
  Matches    7;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         17 ALMRRAVSL 25
              ||:||||:|
Db         17 ALIRRAVTL 25


RESULT 7
Q9R4W6
ID   Q9R4W6      PRELIMINARY;      PRT;    20 AA.
AC   Q9R4W6;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   56 kDa major heat shock protein (Fragment).
OS   Helicobacter pylori (Campylobacter pylori).
OC   Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC   Helicobacteraceae; Helicobacter.
OX   NCBI_TaxID=210;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=95020803; PubMed=7935068;
RA   Yokota K., Hirai Y., Haque M., Hayashi S., Isogai H., Sugiyama T.,
RA   Nagamachi E., Tsukada Y., Fujii N., Oguma K.;
RT   "Heat shock protein produced by Helicobacter pylori.";
RL   Microbiol. Immunol. 38:403-405(1994).
SQ   SEQUENCE   20 AA;  2326 MW;  995EEE7C51529BAC CRC64;

  Query Match             22.0%;  Score 31;  DB 2;  Length 20;
  Best Local Similarity   50.0%;  Pred. No. 8.2e+02;
  Matches    5;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         10 PHSLSSEALM 19
              |::|:|| :|
Db         10 PYNLASEVIM 19


RESULT 8
O57547
ID   O57547      PRELIMINARY;      PRT;    27 AA.
AC   O57547;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Homeobox protein LpHox4-7B (Fragment).
OS   Lampetra planeri (Brook lamprey).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
OC   Petromyzontiformes; Petromyzontidae; Lampetra.
OX   NCBI_TaxID=7750;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98358009; PubMed=9694633;
RA   Sharman A.C., Holland P.W.;
RT   "Estimation of Hox gene cluster number in lampreys.";
RL   Int. J. Dev. Biol. 42:617-620(1998).
DR   EMBL; AF044803; AAC03007.1; -.
FT   NON_TER       1      1
FT   NON_TER      27     27
SQ   SEQUENCE   27 AA;  3056 MW;  650B24846C668637 CRC64;

  Query Match             21.3%;  Score 30;  DB 13;  Length 27;
  Best Local Similarity   35.0%;  Pred. No. 1.6e+03;
  Matches    7;  Conservative    2;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          3 PIAQKSEPHSLSSEALMRRA 22
              |:     ||   : ||  ||
Db          4 PLPHPPAPHRDRARALPHRA 23


RESULT 9
Q9T2P9
ID   Q9T2P9      PRELIMINARY;      PRT;    30 AA.
AC   Q9T2P9;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   Heat shock protein 60 (Fragment).
OS   Narcissus pseudonarcissus (Daffodil).
OG   Mitochondrion.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Amaryllidaceae;
OC   Narcissus.
OX   NCBI_TaxID=39639;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=96291727; PubMed=8754688;
RA   Bonk M., Tadros M., Vandekerckhove J., Al-Babili S., Beyer P.;
RT   "Purification and characterization of chaperonin 60 and heat-shock
RT   protein 70 from chromoplasts of Narcissus pseudonarcissus.";
RL   Plant Physiol. 111:931-939(1996).
DR   HSSP; P06139; 1AON.
SQ   SEQUENCE   30 AA;  3233 MW;  AF5AF69899CE2851 CRC64;

  Query Match             21.3%;  Score 30;  DB 8;  Length 30;
  Best Local Similarity   32.0%;  Pred. No. 1.8e+03;
  Matches    8;  Conservative    4;  Mismatches   13;  Indels    0;  Gaps    0;

Qy          5 AQKSEPHSLSSEALMRRAVSLVTDS 29
              | |     : : ||| | |  : |:
Db          1 AAKDIKFGVEARALMLRGVEELADA 25


RESULT 10
Q8WXC8
ID   Q8WXC8      PRELIMINARY;      PRT;    18 AA.
AC   Q8WXC8;
```

```
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Neuronal nicotinic receptor beta 4 subunit (Fragment).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Valor L.M., Campos-Caro A., Carrasco-Serrano C., Ortiz J.A.,
RA   Ballesta J.J., Criado M.;
RT   "Transcription Factors NF-Y and Sp1 are Important Determinants of the
RT   Promoter Activity of the Bovine and Human Neuronal Nicotinic Receptor
RT   Beta4 Subunit Genes.";
RL   J. Biol. Chem. 0:0-0(2002).
DR   EMBL; AF453877; AAL57840.1; -.
KW   Receptor.
FT   NON_TER      18     18
SQ   SEQUENCE   18 AA;  2050 MW;  69CB11571758876B CRC64;

  Query Match             20.6%;  Score 29;  DB 4;  Length 18;
  Best Local Similarity   87.5%;  Pred. No. 1.5e+03;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         19 MRRAVSLV 26
              |||| ||| 
Db          1 MRRAPSLV 8


RESULT 11
Q8IUA3
ID   Q8IUA3      PRELIMINARY;      PRT;    24 AA.
AC   Q8IUA3;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   ATP-binding cassette subfamily C member 4 variant 1 (ATP-binding
DE   cassette subfamily C member 4 variant 3).
GN   ABCC4.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   PubMed=12499391;
RA   Lamba J.K., Adachi M., Sun D., Tammur J., Schuetz E.G., Allikmets R.,
RA   Schuetz J.D.;
RT   "Nonsense mediated decay downregulates conserved alternatively spliced
RT   ABCC4 transcripts bearing nonsense codons.";
RL   Hum. Mol. Genet. 12:99-109(2003).
DR   EMBL; AY133679; AAN08629.1; -.
DR   EMBL; AY133680; AAN08630.1; -.
KW   ATP-binding.
SQ   SEQUENCE   24 AA;  2809 MW;  12D5E2BB58D93078 CRC64;

  Query Match             20.6%;  Score 29;  DB 4;  Length 24;
  Best Local Similarity   30.0%;  Pred. No. 2.1e+03;
  Matches    6;  Conservative    5;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          2 VPIAQKSEPHSLSSEALMRR 21
              :|: |: :|: |    |  |
Db          2 LPVYQEVKPNPLQDANLCSR 21


RESULT 12
Q9GB51
ID   Q9GB51      PRELIMINARY;      PRT;    27 AA.
AC   Q9GB51;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Cytochrome c oxidase subunit II (Fragment).
GN   COII.
OS   Zosterops nigrorum (yellowish white-eye).
OG   Mitochondrion.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Passeriformes; Zosteropidae; Zosterops.
OX   NCBI_TaxID=135985;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Slikas B., Jones I.B., Derrickson S.R., Fleischer R.C.;
RT   "Phylogenetic relationships of Micronesian white-eyes (Zosteropidae)
RT   based on mitochondrial sequence data.";
RL   Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF168437; AAG12279.1; -.
DR   InterPro; IPR002429; Cyt_c_ox_2.
DR   Pfam; PF00116; COX2; 1.
KW   Oxidoreductase; Mitochondrion.
FT   NON_TER      1      1
SQ   SEQUENCE   27 AA;  2907 MW;  873A2142A87932E5 CRC64;

  Query Match             20.6%;  Score 29;  DB 8;  Length 27;
  Best Local Similarity   53.3%;  Pred. No. 2.4e+03;
  Matches    8;  Conservative    1;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 AVPIAQKSEPHSLSS 15
              |||:|      ||||
Db         12 AVPLANFESWSSLSS 26


RESULT 13
Q9GB48
ID   Q9GB48      PRELIMINARY;      PRT;    27 AA.
AC   Q9GB48;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Cytochrome c oxidase subunit II (Fragment).
GN   COII.
OS   Zosterops montanus (Mountain white-eye).
OG   Mitochondrion.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Passeriformes; Zosteropidae; Zosterops.
OX   NCBI_TaxID=135984;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Slikas B., Jones I.B., Derrickson S.R., Fleischer R.C.;
RT   "Phylogenetic relationships of Micronesian white-eyes (Zosteropidae)
RT   based on mitochondrial sequence data.";
RL   Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF168441; AAG12284.1; -.
DR   InterPro; IPR002429; Cyt_c_ox_2.
DR   Pfam; PF00116; COX2; 1.
KW   Oxidoreductase; Mitochondrion.
FT   NON_TER      1      1
SQ   SEQUENCE   27 AA;  2907 MW;  873A2142A87932E5 CRC64;

  Query Match             20.6%;  Score 29;  DB 8;  Length 27;
  Best Local Similarity   53.3%;  Pred. No. 2.4e+03;
  Matches    8;  Conservative    1;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 AVPIAQKSEPHSLSS 15
              |||:|      ||||
Db         12 AVPLANFESWSSLSS 26


RESULT 14
Q9GB55
ID   Q9GB55      PRELIMINARY;      PRT;    27 AA.
AC   Q9GB55;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Cytochrome c oxidase subunit II (Fragment).
```

```
GN   COII.
OS   Zosterops semperi (Caroline white-eye).
OG   Mitochondrion.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Passeriformes; Zosteropidae; Zosterops.
OX   NCBI_TaxID=135988;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Slikas B., Jones I.B., Derrickson S.R., Fleischer R.C.;
RT   "Phylogenetic relationships of Micronesian white-eyes (Zosteropidae)
RT   based on mitochondrial sequence data.";
RL   Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF168434; AAG12275.1; -.
DR   InterPro; IPR002429; Cyt_c_ox_2.
DR   Pfam; PF00116; COX2; 1.
KW   Oxidoreductase; Mitochondrion.
FT   NON_TER       1      1
SQ   SEQUENCE   27 AA;  2907 MW;  873A2142A87932E5 CRC64;

  Query Match             20.6%;  Score 29;  DB 8;  Length 27;
  Best Local Similarity   53.3%;  Pred. No. 2.4e+03;
  Matches    8;  Conservative    1;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 AVPIAQKSEPHSLSS 15
              |||:|      ||||
Db         12 AVPLANFESWSSLSS 26


RESULT 15
Q9GB45
ID   Q9GB45      PRELIMINARY;      PRT;    27 AA.
AC   Q9GB45;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Cytochrome c oxidase subunit II (Fragment).
GN   COII.
OS   Zosterops poliogaster poliogaster (Heuglin's white-eye).
OG   Mitochondrion.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Passeriformes; Zosteropidae; Zosterops.
OX   NCBI_TaxID=135994;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Slikas B., Jones I.B., Derrickson S.R., Fleischer R.C.;
RT   "Phylogenetic relationships of Micronesian white-eyes (Zosteropidae)
RT   based on mitochondrial sequence data.";
RL   Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF168443; AAG12287.1; -.
DR   InterPro; IPR002429; Cyt_c_ox_2.
DR   Pfam; PF00116; COX2; 1.
KW   Oxidoreductase; Mitochondrion.
FT   NON_TER       1      1
SQ   SEQUENCE   27 AA;  2907 MW;  873A2142A87932E5 CRC64;

  Query Match             20.6%;  Score 29;  DB 8;  Length 27;
  Best Local Similarity   53.3%;  Pred. No. 2.4e+03;
  Matches    8;  Conservative    1;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 AVPIAQKSEPHSLSS 15
              |||:|      ||||
Db         12 AVPLANFESWSSLSS 26


Search completed: October  2, 2003, 09:42:55
Job time : 33 secs
```

GenCore version 5.1.6


OM protein - protein search, using sw model

Run on: October 2, 2003, 09:41:14 ; Search time 16 Seconds
(without alignments)
79.333 Million cell updates/sec

Title: US-09-939-293A-19_COPY_56_85
Perfect score: 141
Sequence: 1 AVPIAQKSEPHSLSSEALMRRAVSLVTDST 30

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 180107

Minimum DB seq length: 0
Maximum DB seq length: 30

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 33.5 | 23.8 | 26 | 3 | US-08-755-587-216 | Sequence 216, App |
| 2 | 33 | 23.4 | 30 | 1 | US-08-031-148-6 | Sequence 6, Appli |
| 3 | 33 | 23.4 | 30 | 2 | US-08-306-078-3 | Sequence 3, Appli |
| 4 | 33 | 23.4 | 30 | 3 | US-08-415-838-6 | Sequence 6, Appli |
| 5 | 33 | 23.4 | 30 | 4 | US-09-354-231B-29 | Sequence 29, Appl |
| 6 | 33 | 23.4 | 30 | 4 | US-09-205-169-6 | Sequence 6, Appli |
| 7 | 33 | 23.4 | 30 | 4 | US-09-128-602B-29 | Sequence 29, Appl |
| 8 | 32 | 22.7 | 24 | 2 | US-08-997-080-102 | Sequence 102, App |
| 9 | 32 | 22.7 | 24 | 2 | US-08-997-362-102 | Sequence 102, App |
| 10 | 32 | 22.7 | 24 | 3 | US-08-873-970-102 | Sequence 102, App |
| 11 | 32 | 22.7 | 24 | 3 | US-09-095-855-102 | Sequence 102, App |
| 12 | 32 | 22.7 | 24 | 4 | US-09-324-542-102 | Sequence 102, App |
| 13 | 32 | 22.7 | 24 | 4 | US-09-205-426-102 | Sequence 102, App |
| 14 | 30.5 | 21.6 | 29 | 2 | US-08-538-711A-16 | Sequence 16, Appl |
| 15 | 30.5 | 21.6 | 29 | 3 | US-08-725-027-16 | Sequence 16, Appl |
| 16 | 30.5 | 21.6 | 29 | 4 | US-09-542-552-16 | Sequence 16, Appl |
| 17 | 30 | 21.3 | 29 | 4 | US-08-469-260A-254 | Sequence 254, App |
| 18 | 30 | 21.3 | 29 | 4 | US-08-488-446-254 | Sequence 254, App |
| 19 | 30 | 21.3 | 29 | 4 | US-08-467-344A-254 | Sequence 254, App |
| 20 | 29 | 20.6 | 12 | 4 | US-09-039-642B-5 | Sequence 5, Appli |
| 21 | 29 | 20.6 | 13 | 4 | US-10-053-485-3 | Sequence 3, Appli |
| 22 | 29 | 20.6 | 20 | 4 | US-09-579-664B-27 | Sequence 27, Appl |
| 23 | 29 | 20.6 | 21 | 3 | US-08-746-111-11 | Sequence 11, Appl |
| 24 | 29 | 20.6 | 23 | 2 | US-08-031-538-57 | Sequence 57, Appl |
| 25 | 29 | 20.6 | 23 | 2 | US-08-031-538-62 | Sequence 62, Appl |
| 26 | 29 | 20.6 | 26 | 3 | US-08-433-522A-47 | Sequence 47, Appl |
| 27 | 29 | 20.6 | 26 | 3 | US-09-135-166-47 | Sequence 47, Appl |
| 28 | 29 | 20.6 | 26 | 3 | US-08-942-046-47 | Sequence 47, Appl |
| 29 | 29 | 20.6 | 29 | 4 | US-09-314-268-123 | Sequence 123, App |
| 30 | 29 | 20.6 | 30 | 3 | US-08-746-411A-14 | Sequence 14, Appl |
| 31 | 29 | 20.6 | 30 | 4 | US-08-857-046A-14 | Sequence 14, Appl |
| 32 | 29 | 20.6 | 30 | 4 | US-09-573-252-14 | Sequence 14, Appl |
| 33 | 28.5 | 20.2 | 21 | 1 | US-08-279-058B-56 | Sequence 56, Appl |
| 34 | 28.5 | 20.2 | 21 | 4 | US-08-828-323-56 | Sequence 56, Appl |
| 35 | 28.5 | 20.2 | 25 | 4 | US-09-205-258-806 | Sequence 806, App |
| 36 | 28 | 19.9 | 19 | 2 | US-08-584-671-5 | Sequence 5, Appli |
| 37 | 28 | 19.9 | 19 | 3 | US-09-027-376-5 | Sequence 5, Appli |
| 38 | 28 | 19.9 | 19 | 3 | US-09-094-192-5 | Sequence 5, Appli |
| 39 | 28 | 19.9 | 21 | 4 | US-09-387-418A-36 | Sequence 36, Appl |
| 40 | 28 | 19.9 | 29 | 4 | US-09-227-357-668 | Sequence 668, App |
| 41 | 27.5 | 19.5 | 29 | 4 | US-09-288-143-150 | Sequence 150, App |
| 42 | 27 | 19.1 | 13 | 4 | US-09-752-165-43 | Sequence 43, Appl |
| 43 | 27 | 19.1 | 20 | 4 | US-09-400-564-21 | Sequence 21, Appl |
| 44 | 27 | 19.1 | 21 | 2 | US-08-832-877-3 | Sequence 3, Appli |
| 45 | 27 | 19.1 | 21 | 4 | US-09-177-165A-46 | Sequence 46, Appl |

ALIGNMENTS

RESULT 1
US-08-755-587-216
; Sequence 216, Application US/08755587
; Patent No. 6045997
; GENERAL INFORMATION:
; APPLICANT: Futreal, Phillip A
; APPLICANT: Wooster, Richard F
; APPLICANT: Ashworth, Alan
; APPLICANT: Stratton, Michael R
; TITLE OF INVENTION: Materials and methods relating to the
; TITLE OF INVENTION: identification and sequencing of the BRCA2 cancer
; TITLE OF INVENTION: susceptibility gene and uses thereof.
; NUMBER OF SEQUENCES: 222
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bell Seltzer Park & Gibson
; STREET: 310 UCB Plaza, 3605 Glenwood Avenue, PO Drawer 31107
; CITY: Raleigh
; STATE: NC
; COUNTRY: USA
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/755,587
; FILING DATE: 25-NOV-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9523959.6
; FILING DATE: 23-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9525555.0
; FILING DATE: 14-DEC-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9617961.9
; FILING DATE: 28-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Kenneth D Sibley
; REGISTRATION NUMBER: 31,665
; REFERENCE/DOCKET NUMBER: 5405-135
; INFORMATION FOR SEQ ID NO: 216:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-755-587-216

Query Match 23.8%; Score 33.5; DB 3; Length 26;
Best Local Similarity 52.9%; Pred. No. 59;
Matches 9; Conservative 5; Mismatches 2; Indels 1; Gaps 1;

```
Qy          12 SLSSEALMRRAVSLVTD 28
               ::||||| ::|| | :|
Db          11 NVSSEAL-QKAVKLFSD 26


RESULT 2
US-08-031-148-6
; Sequence 6, Application US/08031148
; Patent No. 5424398
;  GENERAL INFORMATION:
;    APPLICANT:  Middeldorp, Jaap Michiel.
;    TITLE OF INVENTION:  Peptides and nucleic acid sequences
;    TITLE OF INVENTION:  related to the Epstein-Barr virus.
;    NUMBER OF SEQUENCES:  22
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Akzo Pharma
;      STREET:  1330-A Piccard Drive
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20850-4377
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  Patentin Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/031,148
;      FILING DATE:  19930312
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  EP 92200721.6
;      FILING DATE:  13-MAR-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bobrowicz, Donna
;      REGISTRATION NUMBER:  32,196
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  30 amino acids
;      TYPE:  AMINO ACID
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    ORIGINAL SOURCE:
;      ORGANISM:  Epstein-Barr virus
US-08-031-148-6

  Query Match             23.4%;  Score 33;  DB 1;  Length 30;
  Best Local Similarity   37.5%;  Pred. No. 86;
  Matches    9;  Conservative    4;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          4 IAQKSEPHSLSSEALMRRAVSLVT 27
              :|| : |   || : :| | |  |
Db          4 VAQSATPSVSSSISSLRAATSGAT 27


RESULT 3
US-08-306-078-3
; Sequence 3, Application US/08306078
; Patent No. 5827646
;  GENERAL INFORMATION:
;    APPLICANT:  Middeldorp, Jaap Michiel JM
;    APPLICANT:  van Grunsven, Wouterus Marinus Johannes WMJ
;    TITLE OF INVENTION:  Diagnostic reagents for the
;    TITLE OF INVENTION:  detection of antibodies to EBV.
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  AKZO NOBEL PHARMA
;      STREET:  1330 Piccard Drive
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20850-4377
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/306,078
;      FILING DATE:  14-SEP-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  EP 93202659.4
;      FILING DATE:  14-SEP-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Blackstone, William B.
;      REGISTRATION NUMBER:  29,772
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  30 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-306-078-3

  Query Match             23.4%;  Score 33;  DB 2;  Length 30;
  Best Local Similarity   37.5%;  Pred. No. 86;
  Matches    9;  Conservative    4;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          4 IAQKSEPHSLSSEALMRRAVSLVT 27
              :|| : |   || : :| | |  |
Db          4 VAQSATPSVSSSISSLRAATSGAT 27


RESULT 4
US-08-415-838-6
; Sequence 6, Application US/08415838
; Patent No. 6008327
;  GENERAL INFORMATION:
;    APPLICANT:  Middeldorp, Jaap Michiel.
;    TITLE OF INVENTION:  Peptides and nucleic acid sequences
;    TITLE OF INVENTION:  related to the Epstein-Barr virus.
;    NUMBER OF SEQUENCES:  22
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Akzo-No. 6008327el Patent Department
;      STREET:  1300 Piccard Drive, Suite 206
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  Patentin Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/415,838
;      FILING DATE:  03-APR-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  EP 92200721.6
;      FILING DATE:  13-MAR-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Gormley, Mary E.
;      REGISTRATION NUMBER:  34,409
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  30 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    ORIGINAL SOURCE:
;      ORGANISM:  Epstein-Barr virus
```

```
US-08-415-838-6

  Query Match             23.4%;  Score 33;  DB 3;  Length 30;
  Best Local Similarity   37.5%;  Pred. No. 86;
  Matches    9;  Conservative    4;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          4 IAQKSEPHSLSSEALMRRAVSLVT 27
              :|| : |   || : :| | |  |
Db          4 VAQSATPSVSSSISSLRAATSGAT 27


RESULT 5
US-09-354-231B-29
; Sequence 29, Application US/09354231B
; Patent No. 6342658
; GENERAL INFORMATION:
;  APPLICANT: DeBonte, Lorin R.
;  APPLICANT:  Shorrosh, Basil S.
;  TITLE OF INVENTION: FATTY ACID DESATURASES AND MUTANT SEQUENCES THEREOF
;  FILE REFERENCE: 07148-063002
;  CURRENT APPLICATION NUMBER: US/09/354,231B
;  CURRENT FILING DATE:  1999-07-16
;  PRIOR APPLICATION NUMBER: US 08/874,109
;  PRIOR FILING DATE: 1997-06-12
;  NUMBER OF SEQ ID NOS: 69
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 29
;   LENGTH: 30
;   TYPE: PRT
;   ORGANISM: Zea mays
US-09-354-231B-29

  Query Match             23.4%;  Score 33;  DB 4;  Length 30;
  Best Local Similarity   33.3%;  Pred. No. 86;
  Matches    9;  Conservative    5;  Mismatches   13;  Indels    0;  Gaps    0;

Qy          4 IAQKSEPHSLSSEALMRRAVSLVTDST 30
              || :   |: |  :|:   | ||  |:
Db          4 IAHECGHHAFSDYSLLDDVVGLVLHSS 30


RESULT 6
US-09-205-169-6
; Sequence 6, Application US/09205169
; Patent No. 6365717
;   GENERAL INFORMATION:
;        APPLICANT: Middeldorp, Jaap Michiel.
;        TITLE OF INVENTION: Peptides and nucleic acid sequences
;                            related to the Epstein-Barr virus.
;        NUMBER OF SEQUENCES: 22
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Akzo-No. 6365717el Patent Department
;             STREET: 1300 Piccard Drive, Suite 206
;             CITY: Rockville
;             STATE: Maryland
;             COUNTRY: USA
;             ZIP: 20850
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: Patentin Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/205,169
;             FILING DATE: 04-Dec-1998
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/415,838
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Gormley, Mary E.
;             REGISTRATION NUMBER: 34,409
;   INFORMATION FOR SEQ ID NO: 6:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 30 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: peptide
;        ORIGINAL SOURCE:
;             ORGANISM: Epstein-Barr virus
;        SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-205-169-6

  Query Match             23.4%;  Score 33;  DB 4;  Length 30;
  Best Local Similarity   37.5%;  Pred. No. 86;
  Matches    9;  Conservative    4;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          4 IAQKSEPHSLSSEALMRRAVSLVT 27
              :|| : |   || : :| | |  |
Db          4 VAQSATPSVSSSISSLRAATSGAT 27


RESULT 7
US-09-128-602B-29
; Sequence 29, Application US/09128602B
; Patent No. 6414223
; GENERAL INFORMATION:
;  APPLICANT: Kodali, Dharma
;  APPLICANT:  Fan, Zhegong
;  APPLICANT:  DeBonte, Lorin R.
;  TITLE OF INVENTION: PLANTS, SEEDS AND OILS HAVING AN ELEVATED TOTAL MONOUNSATURATH
;  TITLE OF INVENTION:  FATTY ACID CONTENT
;  FILE REFERENCE: 07148-072001
;  CURRENT APPLICATION NUMBER: US/09/128,602B
;  CURRENT FILING DATE:  1998-08-03
;  NUMBER OF SEQ ID NOS: 68
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 29
;   LENGTH: 30
;   TYPE: PRT
;   ORGANISM: Zea mays
US-09-128-602B-29

  Query Match             23.4%;  Score 33;  DB 4;  Length 30;
  Best Local Similarity   33.3%;  Pred. No. 86;
  Matches    9;  Conservative    5;  Mismatches   13;  Indels    0;  Gaps    0;

Qy          4 IAQKSEPHSLSSEALMRRAVSLVTDST 30
              || :   |: |  :|:   | ||  |:
Db          4 IAHECGHHAFSDYSLLDDVVGLVLHSS 30


RESULT 8
US-08-997-080-102
; Sequence 102, Application US/08997080
; Patent No. 5968524
;  GENERAL INFORMATION:
;    APPLICANT:  WATSON, JAMES D.
;    APPLICANT:  TAN, PAUL L.J.
;    TITLE OF INVENTION:  METHODS AND COMPOUNDS FOR THE TREATMENT OF IMMUNOLOGICALLY
;    NUMBER OF SEQUENCES:  194
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Law Offices of Ann W. Speckman
;      STREET:  2601 Elliott Avenue, Suite 4185
;      CITY:  Seattle
;      STATE:  WA
;      COUNTRY:  USA
;      ZIP:  98121
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/997,080
;      FILING DATE:
```

```
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sleath, Janet
;      REGISTRATION NUMBER:  37,007
;      REFERENCE/DOCKET NUMBER:  11000.1007
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  206-269-0565
;      TELEFAX:  206-269-0563
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  102:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  24 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-997-080-102

  Query Match             22.7%;  Score 32;  DB 2;  Length 24;
  Best Local Similarity   46.7%;  Pred. No. 94;
  Matches    7;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy         15 SEALMRRAVSLVTDS 29
              :|  | :|||:  ||
Db          9 AEEKMEKAVSVARDS 23


RESULT 9
US-08-997-362-102
; Sequence 102, Application US/08997362
; Patent No. 5985287
;  GENERAL INFORMATION:
;    APPLICANT:  Tan, Paul
;    APPLICANT:  Hiyama, Jun
;    APPLICANT:  Visser, Elizabeth
;    APPLICANT:  Skinner, Margot
;    APPLICANT:  Scott, Linda
;    APPLICANT:  Prestidge, Ross
;    TITLE OF INVENTION:  COMPOUNDS AND METHODS FOR
;    TITLE OF INVENTION:  TREATMENT AND DIAGNOSIS OF MYCOBACTERIAL INFECTIONS
;    NUMBER OF SEQUENCES:  194
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Law Offices of Ann W. Speckman
;      STREET:  2601 Elliott Avenue, Suite 4185
;      CITY:  Seattle
;      STATE:  WA
;      COUNTRY:  USA
;      ZIP:  98121
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/997,362
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  U.S. Patent Application No. 5985287 08/873,970
;      FILING DATE:  June 12, 1997
;      APPLICATION NUMBER:  U.S. Patent Application No. 5985287 08/705,347
;      FILING DATE:  August 29, 1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sleath, Janet
;      REGISTRATION NUMBER:  37,007
;      REFERENCE/DOCKET NUMBER:  11000.1002c2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  206-269-0565
;      TELEFAX:  206-269-0563
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  102:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  24 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-997-362-102

  Query Match             22.7%;  Score 32;  DB 2;  Length 24;
  Best Local Similarity   46.7%;  Pred. No. 94;
  Matches    7;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy         15 SEALMRRAVSLVTDS 29
              :|  | :|||:  ||
Db          9 AEEKMEKAVSVARDS 23


RESULT 10
US-08-873-970-102
; Sequence 102, Application US/08873970
; Patent No. 6001361
;  GENERAL INFORMATION:
;    APPLICANT:  Tan, Paul
;    APPLICANT:  Hiyama, Jun
;    APPLICANT:  Visser, Elizabeth
;    APPLICANT:  Skinner, Margot
;    APPLICANT:  Scott, Linda
;    APPLICANT:  Prestidge, Ross
;    TITLE OF INVENTION:  COMPOUNDS AND METHODS FOR
;    TITLE OF INVENTION:  TREATMENT AND DIAGNOSIS OF MYCOBACTERIAL INFECTIONS
;    NUMBER OF SEQUENCES:  106
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Law Offices of Ann W. Speckman
;      STREET:  2601 Elliott Avenue, Suite 4185
;      CITY:  Seattle
;      STATE:  WA
;      COUNTRY:  USA
;      ZIP:  98121
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/873,970
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/705,347
;      FILING DATE:  29-AUG-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sleath, Janet
;      REGISTRATION NUMBER:  37,007
;      REFERENCE/DOCKET NUMBER:  11000.1002C1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  206-269-0565
;      TELEFAX:  206-269-0563
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  102:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  24 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-873-970-102

  Query Match             22.7%;  Score 32;  DB 3;  Length 24;
  Best Local Similarity   46.7%;  Pred. No. 94;
  Matches    7;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy         15 SEALMRRAVSLVTDS 29
```

```
              :|  | :|||:  ||
Db          9 AEEKMEKAVSVARDS 23


RESULT 11
US-09-095-855-102
; Sequence 102, Application US/09095855
; Patent No. 6160093
;  GENERAL INFORMATION:
;    APPLICANT:  Tan, Paul
;    APPLICANT:  Visser, Elizabeth
;    APPLICANT:  Skinner, Margot
;    APPLICANT:  Prestidge, Ross
;    TITLE OF INVENTION:  Compounds and Methods for
;    TITLE OF INVENTION:  Treatment and Diagnosis of Mycobacterial Infections
;    NUMBER OF SEQUENCES:  208
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Law Offices of Ann W. Speckman
;      STREET:  2601 Elliott Avenue, Suite 4185
;      CITY:  Seattle
;      STATE:  WA
;      COUNTRY:  USA
;      ZIP:  98121
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/095,855
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/705,347
;      FILING DATE:  29-AUG-1996
;      APPLICATION NUMBER:  08/873,970
;      FILING DATE:  12-JUN-1997
;      APPLICATION NUMBER:  08/997,362
;      FILING DATE:  23-DEC-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sleath, Janet
;      REGISTRATION NUMBER:  37,007
;      REFERENCE/DOCKET NUMBER:  11000.1002c3
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  206-269-0565
;      TELEFAX:  206-269-0563
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  102:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  24 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-095-855-102

  Query Match             22.7%;  Score 32;  DB 3;  Length 24;
  Best Local Similarity   46.7%;  Pred. No. 94;
  Matches    7;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy         15 SEALMRRAVSLVTDS 29
              :|  | :|||:  ||
Db          9 AEEKMEKAVSVARDS 23


RESULT 12
US-09-324-542-102
; Sequence 102, Application US/09324542
; Patent No. 6328978
; GENERAL INFORMATION:
;  APPLICANT: Watson, James D.
;  APPLICANT: Tan, Paul L.J.
;  APPLICANT: Prestidge, Ross
;  TITLE OF INVENTION: Methods and Compounds for the Treatment
;  TITLE OF INVENTION: of Immunologically-Mediated Skin Disorders
;  FILE REFERENCE: 11000.1007c1
;  CURRENT APPLICATION NUMBER: US/09/324,542
;  CURRENT FILING DATE: 1999-06-02
;  EARLIER APPLICATION NUMBER: US 08/997,080
;  EARLIER FILING DATE: 1997-12-23
;  NUMBER OF SEQ ID NOS: 194
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 102
;   LENGTH: 24
;   TYPE: PRT
;   ORGANISM: Mycobacterium vaccae
;   FEATURE:
;   NAME/KEY: UNSURE
;   LOCATION: (1)...(1)
US-09-324-542-102

  Query Match             22.7%;  Score 32;  DB 4;  Length 24;
  Best Local Similarity   46.7%;  Pred. No. 94;
  Matches    7;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy         15 SEALMRRAVSLVTDS 29
              :|  | :|||:  ||
Db          9 AEEKMEKAVSVARDS 23


RESULT 13
US-09-205-426-102
; Sequence 102, Application US/09205426
; Patent No. 6406704
; GENERAL INFORMATION:
;  APPLICANT: Watson, James D.
;  APPLICANT: Tan, Paul L. J.
;  TITLE OF INVENTION: Compounds and Methods for Treatment and
;  TITLE OF INVENTION: Diagnosis of Mycobacterial Infections
;  FILE REFERENCE: 11000.1002c4
;  CURRENT APPLICATION NUMBER: US/09/205,426
;  CURRENT FILING DATE: 1998-12-04
;  EARLIER APPLICATION NUMBER: 09/095,855
;  EARLIER FILING DATE: 1998-06-11
;  EARLIER APPLICATION NUMBER: 08/997,362
;  EARLIER FILING DATE: 1997-12-23
;  EARLIER APPLICATION NUMBER: 08/873,970
;  EARLIER FILING DATE: 1997-06-12
;  EARLIER APPLICATION NUMBER: 08/705,347
;  EARLIER FILING DATE: 1996-08-29
;  NUMBER OF SEQ ID NOS: 208
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 102
;   LENGTH: 24
;   TYPE: PRT
;   ORGANISM: Mycobacterium vaccae
;   FEATURE:
;   NAME/KEY: UNSURE
;   LOCATION: (1)...(1)
US-09-205-426-102

  Query Match             22.7%;  Score 32;  DB 4;  Length 24;
  Best Local Similarity   46.7%;  Pred. No. 94;
  Matches    7;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy         15 SEALMRRAVSLVTDS 29
              :|  | :|||:  ||
Db          9 AEEKMEKAVSVARDS 23


RESULT 14
US-08-538-711A-16
; Sequence 16, Application US/08538711A
; Patent No. 5994062
;  GENERAL INFORMATION:
;    APPLICANT:  MULSHINE, JAMES, L.
```

```
;    TITLE OF INVENTION:  AN EPITHELIAL PROTEIN AND
;    TITLE OF INVENTION:  DNA THEREOF FOR USE IN EARLY CANCER DETECTION
;    NUMBER OF SEQUENCES:  23
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORGAN & FINNEGAN, L.L.P.
;      STREET:  345 PARK AVENUE
;      CITY:  NEW YORK
;      STATE:  NEW YORK
;      COUNTRY:  USA
;      ZIP:  10154
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  FLOPPY DISK
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  ASCII
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/538,711A
;      FILING DATE:  02-OCT-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  KATHRYN M. BROWN
;      REGISTRATION NUMBER:  34,556
;      REFERENCE/DOCKET NUMBER:  2026-4201
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 758-4800
;      TELEFAX:  (212) 751-6849
;  INFORMATION FOR SEQ ID NO:  16:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  29
;      TYPE:  Amino Acid
;      STRANDEDNESS:  Unknown
;      TOPOLOGY:  Linear
;    MOLECULE TYPE:  peptide
US-08-538-711A-16

  Query Match             21.6%;  Score 30.5;  DB 2;  Length 29;
  Best Local Similarity   33.3%;  Pred. No. 2.1e+02;
  Matches    8;  Conservative    4;  Mismatches   11;  Indels    1;  Gaps    1;

Qy          2 VPIAQKSEPHSLSSEAL-MRRAVS 24
              |  |  : || :    :  :|||| 
Db          6 VDAAMNARPHKVDGRVVEPKRAVS 29


RESULT 15
US-08-725-027-16
; Sequence 16, Application US/08725027
; Patent No. 6251586
;  GENERAL INFORMATION:
;    APPLICANT:  MULSHINE, JAMES, L.
;    APPLICANT:  TOCKMAN, MELVYN, S.
;    TITLE OF INVENTION:  AN EPITHELIAL PROTEIN AND
;    TITLE OF INVENTION:  DNA THEREOF FOR USE IN EARLY CANCER DETECTION
;    NUMBER OF SEQUENCES:  23
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORGAN & FINNEGAN, L.L.P.
;      STREET:  345 PARK AVENUE
;      CITY:  NEW YORK
;      STATE:  NEW YORK
;      COUNTRY:  USA
;      ZIP:  10154
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  FLOPPY DISK
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  ASCII
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/725,027
;      FILING DATE:  02-OCT-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US08/538,711
;      FILING DATE:  02-OCT-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  KATHRYN M. BROWN
;      REGISTRATION NUMBER:  34,556
;      REFERENCE/DOCKET NUMBER:  2026-4201US1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 758-4800
;      TELEFAX:  (212) 751-6849
;  INFORMATION FOR SEQ ID NO:  16:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  29
;      TYPE:  Amino Acid
;      STRANDEDNESS:  Unknown
;      TOPOLOGY:  Linear
;    MOLECULE TYPE:  peptide
US-08-725-027-16

  Query Match             21.6%;  Score 30.5;  DB 3;  Length 29;
  Best Local Similarity   33.3%;  Pred. No. 2.1e+02;
  Matches    8;  Conservative    4;  Mismatches   11;  Indels    1;  Gaps    1;

Qy          2 VPIAQKSEPHSLSSEAL-MRRAVS 24
              |  |  : || :    :  :|||| 
Db          6 VDAAMNARPHKVDGRVVEPKRAVS 29


Search completed: October  2, 2003, 09:43:45
Job time : 17 secs
```

GenCore version 5.1.6


OM protein - protein search, using sw model

Run on: October 2, 2003, 09:41:59 ; Search time 23 Seconds
(without alignments)
206.365 Million cell updates/sec

Title: US-09-939-293A-19_COPY_56_85
Perfect score: 141
Sequence: 1 AVPIAQKSEPHSLSSEALMRRAVSLVTDST 30

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 587654 seqs, 158212981 residues

Total number of hits satisfying chosen parameters: 142344

Minimum DB seq length: 0
Maximum DB seq length: 30

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 141 | 100.0 | 30 | 10 | US-09-939-293-7 | Sequence 7, Appli |
| 2 | 70 | 49.6 | 15 | 14 | US-10-068-569-8 | Sequence 8, Appli |
| 3 | 70 | 49.6 | 15 | 15 | US-10-197-634-8 | Sequence 8, Appli |
| 4 | 63 | 44.7 | 13 | 10 | US-09-798-116-20 | Sequence 20, Appl |
| 5 | 63 | 44.7 | 13 | 10 | US-09-798-116-22 | Sequence 22, Appl |
| 6 | 49 | 34.8 | 10 | 10 | US-09-965-967-18 | Sequence 18, Appl |
| 7 | 49 | 34.8 | 13 | 10 | US-09-965-967-25 | Sequence 25, Appl |
| 8 | 40 | 28.4 | 29 | 10 | US-09-798-116-25 | Sequence 25, Appl |
| 9 | 37 | 26.2 | 25 | 9 | US-09-864-761-47499 | Sequence 47499, A |
| 10 | 34 | 24.1 | 25 | 12 | US-10-269-806-80 | Sequence 80, Appl |
| 11 | 33.5 | 23.8 | 29 | 9 | US-09-864-761-40008 | Sequence 40008, A |
| 12 | 33 | 23.4 | 7 | 10 | US-09-939-293-6 | Sequence 6, Appli |
| 13 | 33 | 23.4 | 7 | 10 | US-09-965-967-8 | Sequence 8, Appli |
| 14 | 33 | 23.4 | 7 | 12 | US-10-293-371-1 | Sequence 1, Appli |
| 15 | 33 | 23.4 | 7 | 12 | US-10-293-371-24 | Sequence 24, Appl |
| 16 | 33 | 23.4 | 7 | 12 | US-10-293-371-45 | Sequence 45, Appl |
| 17 | 33 | 23.4 | 7 | 12 | US-10-302-811-5 | Sequence 5, Appli |
| 18 | 33 | 23.4 | 7 | 14 | US-10-068-569-12 | Sequence 12, Appl |
| 19 | 33 | 23.4 | 24 | 11 | US-09-798-889-119 | Sequence 119, App |
| 20 | 33 | 23.4 | 30 | 9 | US-09-995-297-29 | Sequence 29, Appl |
| 21 | 33 | 23.4 | 30 | 12 | US-09-771-904-29 | Sequence 29, Appl |
| 22 | 33 | 23.4 | 30 | 14 | US-10-036-729-6 | Sequence 6, Appli |
| 23 | 32 | 22.7 | 20 | 9 | US-09-864-761-36958 | Sequence 36958, A |
| 24 | 32 | 22.7 | 20 | 15 | US-10-225-567A-1480 | Sequence 1480, Ap |
| 25 | 32 | 22.7 | 21 | 15 | US-10-097-065-542 | Sequence 542, App |
| 26 | 32 | 22.7 | 24 | 11 | US-09-880-505-102 | Sequence 102, App |
| 27 | 32 | 22.7 | 24 | 14 | US-10-051-643-102 | Sequence 102, App |
| 28 | 32 | 22.7 | 29 | 9 | US-09-864-761-40435 | Sequence 40435, A |
| 29 | 30 | 21.3 | 17 | 15 | US-10-225-567A-1084 | Sequence 1084, Ap |
| 30 | 30 | 21.3 | 18 | 9 | US-09-864-761-44657 | Sequence 44657, A |
| 31 | 30 | 21.3 | 23 | 12 | US-09-965-738-279 | Sequence 279, App |
| 32 | 30 | 21.3 | 24 | 9 | US-09-864-761-45203 | Sequence 45203, A |
| 33 | 30 | 21.3 | 25 | 9 | US-09-864-761-34911 | Sequence 34911, A |
| 34 | 30 | 21.3 | 25 | 9 | US-09-864-761-41716 | Sequence 41716, A |
| 35 | 30 | 21.3 | 25 | 9 | US-09-864-761-42933 | Sequence 42933, A |
| 36 | 30 | 21.3 | 27 | 9 | US-09-864-761-40837 | Sequence 40837, A |
| 37 | 30 | 21.3 | 27 | 11 | US-09-899-495-59 | Sequence 59, Appl |
| 38 | 30 | 21.3 | 29 | 8 | US-08-424-550B-254 | Sequence 254, App |
| 39 | 30 | 21.3 | 29 | 9 | US-09-864-761-43946 | Sequence 43946, A |
| 40 | 30 | 21.3 | 30 | 9 | US-09-864-761-42325 | Sequence 42325, A |
| 41 | 29.5 | 20.9 | 27 | 9 | US-09-864-761-40411 | Sequence 40411, A |
| 42 | 29.5 | 20.9 | 29 | 15 | US-10-174-410-272 | Sequence 272, App |
| 43 | 29 | 20.6 | 10 | 15 | US-10-226-007-1533 | Sequence 1533, Ap |
| 44 | 29 | 20.6 | 11 | 15 | US-10-226-007-1534 | Sequence 1534, Ap |
| 45 | 29 | 20.6 | 11 | 15 | US-10-226-007-1547 | Sequence 1547, Ap |

ALIGNMENTS

RESULT 1
US-09-939-293-7
; Sequence 7, Application US/09939293
; Patent No. US20020132786A1
; GENERAL INFORMATION:
; APPLICANT: Alnemri, Emad S.
; TITLE OF INVENTION: AN IAP PEPTIDE OR POLYPEPTIDE
; TITLE OF INVENTION: AND METHODS OF USING THE SAME
; FILE REFERENCE: 480140.465
; CURRENT APPLICATION NUMBER: US/09/939,293
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-939-293-7

Query Match 100.0%; Score 141; DB 10; Length 30;
Best Local Similarity 100.0%; Pred. No. 8.5e-15;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AVPIAQKSEPHSLSSEALMRRAVSLVTDST 30
||||||||||||||||||||||||||||||
Db 1 AVPIAQKSEPHSLSSEALMRRAVSLVTDST 30

RESULT 2
US-10-068-569-8
; Sequence 8, Application US/10068569
; Publication No. US20020160975A1
; GENERAL INFORMATION:
; APPLICANT: Srinivasula, Srinivasa M.
; APPLICANT: Fernandes-Alnemri, Teresa
; APPLICANT: Alnemri, Emad S.
; TITLE OF INVENTION: A CONSERVED XIAP-INTERACTION MOTIF IN
; TITLE OF INVENTION: CASPASE-9 AND SMAC/DIABLO FOR MEDIATING APOPTOSIS

```
;  FILE REFERENCE: 480140.475
;  CURRENT APPLICATION NUMBER: US/10/068,569
;  CURRENT FILING DATE:  2002-02-06
;  NUMBER OF SEQ ID NOS: 28
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8
;   LENGTH: 15
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-068-569-8

  Query Match             49.6%;  Score 70;  DB 14;  Length 15;
  Best Local Similarity   93.3%;  Pred. No. 0.00034;
  Matches   14;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AVPIAQKSEPHSLSS 15
              ||||||||||||||:
Db          1 AVPIAQKSEPHSLSN 15


RESULT 3
US-10-197-634-8
; Sequence 8, Application US/10197634
; Publication No. US20030073629A1
; GENERAL INFORMATION:
;  APPLICANT: Alnemri, Emad S.
;  TITLE OF INVENTION: OMI AND DOMAINS THEREOF THAT DISRUPT
;  TITLE OF INVENTION:  IAP-CASPASE INTERACTION
;  FILE REFERENCE: 480140.479
;  CURRENT APPLICATION NUMBER: US/10/197,634
;  CURRENT FILING DATE:  2002-07-15
;  NUMBER OF SEQ ID NOS: 17
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8
;   LENGTH: 15
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-197-634-8

  Query Match             49.6%;  Score 70;  DB 15;  Length 15;
  Best Local Similarity   93.3%;  Pred. No. 0.00034;
  Matches   14;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AVPIAQKSEPHSLSS 15
              ||||||||||||||:
Db          1 AVPIAQKSEPHSLSN 15


RESULT 4
US-09-798-116-20
; Sequence 20, Application US/09798116
; Patent No. US20020110851A1
; GENERAL INFORMATION:
;  APPLICANT: Verhagen, Anne Marie
;  APPLICANT:  Ekert, Paul
;  APPLICANT:  Vaux, David
;  TITLE OF INVENTION: No. US20020110851A1el Polypeptides, Modulatory Agents Therefor an
;  FILE REFERENCE: 10338-004US
;  CURRENT APPLICATION NUMBER: US/09/798,116
;  CURRENT FILING DATE:  2001-03-02
;  PRIOR APPLICATION NUMBER: AU PQ5995/00
;  PRIOR FILING DATE: 2000-03-02
;  NUMBER OF SEQ ID NOS: 25
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 20
;   LENGTH: 13
;   TYPE: PRT
;   ORGANISM: synthetic
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (12)..(12)
;   OTHER INFORMATION: M is methionine sulfoxide
US-09-798-116-20

  Query Match             44.7%;  Score 63;  DB 10;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 0.0034;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          8 SEPHSLSSEALMR 20
              |||||||||||||
Db          1 SEPHSLSSEALMR 13


RESULT 5
US-09-798-116-22
; Sequence 22, Application US/09798116
; Patent No. US20020110851A1
; GENERAL INFORMATION:
;  APPLICANT: Verhagen, Anne Marie
;  APPLICANT:  Ekert, Paul
;  APPLICANT:  Vaux, David
;  TITLE OF INVENTION: No. US20020110851A1el Polypeptides, Modulatory Agents Therefor
;  FILE REFERENCE: 10338-004US
;  CURRENT APPLICATION NUMBER: US/09/798,116
;  CURRENT FILING DATE:  2001-03-02
;  PRIOR APPLICATION NUMBER: AU PQ5995/00
;  PRIOR FILING DATE: 2000-03-02
;  NUMBER OF SEQ ID NOS: 25
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
;   LENGTH: 13
;   TYPE: PRT
;   ORGANISM: synthetic
US-09-798-116-22

  Query Match             44.7%;  Score 63;  DB 10;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 0.0034;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          8 SEPHSLSSEALMR 20
              |||||||||||||
Db          1 SEPHSLSSEALMR 13


RESULT 6
US-09-965-967-18
; Sequence 18, Application US/09965967
; Patent No. US20020177557A1
; GENERAL INFORMATION:
;  APPLICANT: Shi, Yigong
;  TITLE OF INVENTION: Compositions And Methods For Regulating Apoptosis
;  FILE REFERENCE: PU-0031 (01-1739-1)
;  CURRENT APPLICATION NUMBER: US/09/965,967
;  CURRENT FILING DATE:  2001-09-28
;  PRIOR APPLICATION NUMBER: 60/236,574
;  PRIOR FILING DATE: 2000-09-29
;  PRIOR APPLICATION NUMBER: 60/256,830
;  PRIOR FILING DATE: 2000-12-20
;  NUMBER OF SEQ ID NOS: 30
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
;   LENGTH: 10
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-965-967-18

  Query Match             34.8%;  Score 49;  DB 10;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 0.37;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AVPIAQKSEP 10
              ||||||||||
Db          1 AVPIAQKSEP 10


RESULT 7
```

```
US-09-965-967-25
; Sequence 25, Application US/09965967
; Patent No. US20020177557A1
; GENERAL INFORMATION:
;  APPLICANT: Shi, Yigong
;  TITLE OF INVENTION: Compositions And Methods For Regulating Apoptosis
;  FILE REFERENCE: PU-0031 (01-1739-1)
;  CURRENT APPLICATION NUMBER: US/09/965,967
;  CURRENT FILING DATE:  2001-09-28
;  PRIOR APPLICATION NUMBER: 60/236,574
;  PRIOR FILING DATE: 2000-09-29
;  PRIOR APPLICATION NUMBER: 60/256,830
;  PRIOR FILING DATE: 2000-12-20
;  NUMBER OF SEQ ID NOS: 30
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 25
;   LENGTH: 13
;   TYPE: PRT
;   ORGANISM: Drosophila melanogaster
US-09-965-967-25

  Query Match             34.8%;  Score 49;  DB 10;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 0.5;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AVPIAQKSEP 10
              ||||||||||
Db          4 AVPIAQKSEP 13


RESULT 8
US-09-798-116-25
; Sequence 25, Application US/09798116
; Patent No. US20020110851A1
; GENERAL INFORMATION:
;  APPLICANT: Verhagen, Anne Marie
;  APPLICANT:  Ekert, Paul
;  APPLICANT:  Vaux, David
;  TITLE OF INVENTION: No. US20020110851A1el Polypeptides, Modulatory Agents Therefor an
;  FILE REFERENCE: 10338-004US
;  CURRENT APPLICATION NUMBER: US/09/798,116
;  CURRENT FILING DATE:  2001-03-02
;  PRIOR APPLICATION NUMBER: AU PQ5995/00
;  PRIOR FILING DATE: 2000-03-02
;  NUMBER OF SEQ ID NOS: 25
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 25
;   LENGTH: 29
;   TYPE: PRT
;   ORGANISM: synthetic
US-09-798-116-25

  Query Match             28.4%;  Score 40;  DB 10;  Length 29;
  Best Local Similarity   100.0%;  Pred. No. 31;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         22 AVSLVTDST 30
              |||||||||
Db          1 AVSLVTDST 9


RESULT 9
US-09-864-761-47499
; Sequence 47499, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 47499
;   LENGTH: 25
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC022267.2
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.6
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.7
US-09-864-761-47499

  Query Match             26.2%;  Score 37;  DB 9;  Length 25;
  Best Local Similarity   32.0%;  Pred. No. 76;
  Matches    8;  Conservative    7;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          4 IAQKSEPHSLSSEALMRRAVSLVTD 28
              |::| | |  :   |:   ||:::|
Db          1 ISEKCERHPGTLHPLLGERVSVISD 25


RESULT 10
US-10-269-806-80
; Sequence 80, Application US/10269806
; Publication No. US20030176352A1
; GENERAL INFORMATION:
;  APPLICANT: Min, Hosung
;  APPLICANT:  Sitney, Karen
;  APPLICANT:  Hartley, Cynthia
;  TITLE OF INVENTION: Peptides and Related Compounds Having Thrombopoietic Activity
;  FILE REFERENCE: A-750
;  CURRENT APPLICATION NUMBER: US/10/269,806
;  CURRENT FILING DATE:  2002-10-10
;  NUMBER OF SEQ ID NOS: 199
;  SOFTWARE: PatentIn version 3.1
```

```
; SEQ ID NO 80
;   LENGTH: 25
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthesized Peptide Sequence
US-10-269-806-80

  Query Match             24.1%;  Score 34;  DB 12;  Length 25;
  Best Local Similarity   60.0%;  Pred. No. 2.2e+02;
  Matches    6;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 IAQKSEPHSL 13
              :||:  |||| 
Db         15 LAQRLNPHSL 24


RESULT 11
US-09-864-761-40008
; Sequence 40008, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 40008
;   LENGTH: 29
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC004898.1
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL =   8
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 7.2
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 7.4
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 7.8
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 7.8
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 7.5
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 8.2
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 8.6
;   OTHER INFORMATION: EST_HUMAN HIT: BE144340.1, EVALUE 7.80e+00
US-09-864-761-40008

  Query Match             23.8%;  Score 33.5;  DB 9;  Length 29;
  Best Local Similarity   45.5%;  Pred. No. 3.1e+02;
  Matches   10;  Conservative    2;  Mismatches    7;  Indels    3;  Gaps    1;

Qy          1 AVPIAQKSEP---HSLSSEALM 19
              || : |   |   || || :||
Db          1 AVHLLQPKTPFRWHSKSSHSLM 22


RESULT 12
US-09-939-293-6
; Sequence 6, Application US/09939293
; Patent No. US20020132786A1
; GENERAL INFORMATION:
;  APPLICANT: Alnemri, Emad S.
;  TITLE OF INVENTION: AN IAP PEPTIDE OR POLYPEPTIDE
;  TITLE OF INVENTION:  AND METHODS OF USING THE SAME
;  FILE REFERENCE: 480140.465
;  CURRENT APPLICATION NUMBER: US/09/939,293
;  CURRENT FILING DATE:  2001-08-24
;  NUMBER OF SEQ ID NOS: 18
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
;   LENGTH: 7
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-939-293-6

  Query Match             23.4%;  Score 33;  DB 10;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 5.2e+05;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AVPIAQK 7
              |||||||
Db          1 AVPIAQK 7


RESULT 13
US-09-965-967-8
; Sequence 8, Application US/09965967
; Patent No. US20020177557A1
; GENERAL INFORMATION:
;  APPLICANT: Shi, Yigong
;  TITLE OF INVENTION: Compositions And Methods For Regulating Apoptosis
;  FILE REFERENCE: PU-0031 (01-1739-1)
;  CURRENT APPLICATION NUMBER: US/09/965,967
;  CURRENT FILING DATE:  2001-09-28
;  PRIOR APPLICATION NUMBER: 60/236,574
;  PRIOR FILING DATE: 2000-09-29
;  PRIOR APPLICATION NUMBER: 60/256,830
;  PRIOR FILING DATE: 2000-12-20
;  NUMBER OF SEQ ID NOS: 30
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
;   LENGTH: 7
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-965-967-8

  Query Match             23.4%;  Score 33;  DB 10;  Length 7;
```

Best Local Similarity 100.0%; Pred. No. 5.2e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 AVPIAQK 7
              |||||||
Db          1 AVPIAQK 7
```

RESULT 14
US-10-293-371-1
; Sequence 1, Application US/10293371
; Publication No. US20030157522A1
; GENERAL INFORMATION:
; APPLICANT: BOUDREAULT, ALAIN
; APPLICANT: KORNELUK, ROBERT G.
; APPLICANT: LACASSE, ERIC
; APPLICANT: LISTON, PETER
; TITLE OF INVENTION: Methods and Reagents for Peptide-Bir
; TITLE OF INVENTION: Interaction Screens
; FILE REFERENCE: 07891/030002
; CURRENT APPLICATION NUMBER: US/10/293,371
; CURRENT FILING DATE: 2003-04-08
; PRIOR APPLICATION NUMBER: US 60/370,934
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: US 60/332,300
; PRIOR FILING DATE: 2001-11-09
; NUMBER OF SEQ ID NOS: 85
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-293-371-1

Query Match 23.4%; Score 33; DB 12; Length 7;
Best Local Similarity 100.0%; Pred. No. 5.2e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 AVPIAQK 7
              |||||||
Db          1 AVPIAQK 7
```

RESULT 15
US-10-293-371-24
; Sequence 24, Application US/10293371
; Publication No. US20030157522A1
; GENERAL INFORMATION:
; APPLICANT: BOUDREAULT, ALAIN
; APPLICANT: KORNELUK, ROBERT G.
; APPLICANT: LACASSE, ERIC
; APPLICANT: LISTON, PETER
; TITLE OF INVENTION: Methods and Reagents for Peptide-Bir
; TITLE OF INVENTION: Interaction Screens
; FILE REFERENCE: 07891/030002
; CURRENT APPLICATION NUMBER: US/10/293,371
; CURRENT FILING DATE: 2003-04-08
; PRIOR APPLICATION NUMBER: US 60/370,934
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: US 60/332,300
; PRIOR FILING DATE: 2001-11-09
; NUMBER OF SEQ ID NOS: 85
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-293-371-24

Query Match 23.4%; Score 33; DB 12; Length 7;
Best Local Similarity 100.0%; Pred. No. 5.2e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 AVPIAQK 7
              |||||||
Db          1 AVPIAQK 7
```

Search completed: October 2, 2003, 09:44:16
Job time : 23 secs

```
;  NUMBER OF SEQ ID NOS: 25
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
;   LENGTH: 84
;   TYPE: PRT
;   ORGANISM: Rattus sp.
US-09-798-116-9

  Query Match             88.7%;  Score 125;  DB 10;  Length 84;
  Best Local Similarity   90.0%;  Pred. No. 8.4e-12;
  Matches   27;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 AVPIAQKSEPHSLSSEALMRRAVSLVTDST 30
              |||||||||| |||:|||||||||||||:||
Db         54 AVPIAQKSEPQSLSNEALMRRAVSLVTNST 83


RESULT 12
US-09-798-116-8
; Sequence 8, Application US/09798116
; Patent No. US20020110851A1
; GENERAL INFORMATION:
;  APPLICANT: Verhagen, Anne Marie
;  APPLICANT:  Ekert, Paul
;  APPLICANT:  Vaux, David
;  TITLE OF INVENTION: No. US20020110851A1el Polypeptides, Modulatory Agents Therefor ar
;  FILE REFERENCE: 10338-004US
;  CURRENT APPLICATION NUMBER: US/09/798,116
;  CURRENT FILING DATE:  2001-03-02
;  PRIOR APPLICATION NUMBER: AU PQ5995/00
;  PRIOR FILING DATE: 2000-03-02
;  NUMBER OF SEQ ID NOS: 25
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
;   LENGTH: 177
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-798-116-8

  Query Match             76.6%;  Score 108;  DB 10;  Length 177;
  Best Local Similarity   100.0%;  Pred. No. 8.5e-09;
  Matches   23;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          8 SEPHSLSSEALMRRAVSLVTDST 30
              |||||||||||||||||||||||
Db          1 SEPHSLSSEALMRRAVSLVTDST 23


RESULT 13
US-09-798-116-6
; Sequence 6, Application US/09798116
; Patent No. US20020110851A1
; GENERAL INFORMATION:
;  APPLICANT: Verhagen, Anne Marie
;  APPLICANT:  Ekert, Paul
;  APPLICANT:  Vaux, David
;  TITLE OF INVENTION: No. US20020110851A1el Polypeptides, Modulatory Agents Therefor ar
;  FILE REFERENCE: 10338-004US
;  CURRENT APPLICATION NUMBER: US/09/798,116
;  CURRENT FILING DATE:  2001-03-02
;  PRIOR APPLICATION NUMBER: AU PQ5995/00
;  PRIOR FILING DATE: 2000-03-02
;  NUMBER OF SEQ ID NOS: 25
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
;   LENGTH: 177
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-798-116-6

  Query Match             74.5%;  Score 105;  DB 10;  Length 177;
  Best Local Similarity   95.7%;  Pred. No. 2.5e-08;
  Matches   22;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          8 SEPHSLSSEALMRRAVSLVTDST 30
              |||||||:|||||||||||||||
Db          1 SEPHSLSNEALMRRAVSLVTDST 23


RESULT 14
US-10-068-569-8
; Sequence 8, Application US/10068569
; Publication No. US20020160975A1
; GENERAL INFORMATION:
;  APPLICANT: Srinivasula, Srinivasa M.
;  APPLICANT:  Fernandes-Alnemri, Teresa
;  APPLICANT:  Alnemri, Emad S.
;  TITLE OF INVENTION: A CONSERVED XIAP-INTERACTION MOTIF IN
;  TITLE OF INVENTION:  CASPASE-9 AND SMAC/DIABLO FOR MEDIATING APOPTOSIS
;  FILE REFERENCE: 480140.475
;  CURRENT APPLICATION NUMBER: US/10/068,569
;  CURRENT FILING DATE:  2002-02-06
;  NUMBER OF SEQ ID NOS: 28
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8
;   LENGTH: 15
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-068-569-8

  Query Match             49.6%;  Score 70;  DB 14;  Length 15;
  Best Local Similarity   93.3%;  Pred. No. 0.00034;
  Matches   14;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AVPIAQKSEPHSLSS 15
              |||||||||||||||:
Db          1 AVPIAQKSEPHSLSN 15


RESULT 15
US-10-197-634-8
; Sequence 8, Application US/10197634
; Publication No. US20030073629A1
; GENERAL INFORMATION:
;  APPLICANT: Alnemri, Emad S.
;  TITLE OF INVENTION: OMI AND DOMAINS THEREOF THAT DISRUPT
;  TITLE OF INVENTION:  IAP-CASPASE INTERACTION
;  FILE REFERENCE: 480140.479
;  CURRENT APPLICATION NUMBER: US/10/197,634
;  CURRENT FILING DATE:  2002-07-15
;  NUMBER OF SEQ ID NOS: 17
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8
;   LENGTH: 15
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-197-634-8

  Query Match             49.6%;  Score 70;  DB 15;  Length 15;
  Best Local Similarity   93.3%;  Pred. No. 0.00034;
  Matches   14;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AVPIAQKSEPHSLSS 15
              |||||||||||||||:
Db          1 AVPIAQKSEPHSLSN 15


Search completed: October  2, 2003, 09:41:11
Job time : 24 secs
```

```
;  TITLE OF INVENTION: Methods For Determining the Prognosis
;  TITLE OF INVENTION:  For Cancer Patients Using Tucan
;  FILE REFERENCE: P-LJ 5254
;  CURRENT APPLICATION NUMBER: US/10/141,618
;  CURRENT FILING DATE:  2002-05-07
;  PRIOR APPLICATION NUMBER: US 60/289,233
;  PRIOR FILING DATE: 2001-05-07
;  PRIOR APPLICATION NUMBER: US 60/356,934
;  PRIOR FILING DATE: 2002-02-12
;  PRIOR APPLICATION NUMBER: US 09/388,221
;  PRIOR FILING DATE: 1999-09-01
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14
;   LENGTH: 239
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-141-618-14

  Query Match             100.0%;  Score 141;  DB 12;  Length 239;
  Best Local Similarity   100.0%;  Pred. No. 9.8e-14;
  Matches   30;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AVPIAQKSEPHSLSSEALMRRAVSLVTDST 30
              ||||||||||||||||||||||||||||||
Db         56 AVPIAQKSEPHSLSSEALMRRAVSLVTDST 85


RESULT 8
US-10-153-668-348
; Sequence 348, Application US/10153668
; Publication No. US20030092616A1
; GENERAL INFORMATION:
;  APPLICANT: HONDA, Goichi
;  APPLICANT:  MATSUDA, Akio
;  APPLICANT:  MURAMATSU, Shuji
;  APPLICANT:  ISHIZAWA, Kenya
;  TITLE OF INVENTION: STAT6 Activating Gene
;  FILE REFERENCE: 1254-0207P
;  CURRENT APPLICATION NUMBER: US/10/153,668
;  CURRENT FILING DATE:  2002-05-24
;  PRIOR APPLICATION NUMBER: US 60/293,172
;  PRIOR FILING DATE: 2001-05-25
;  PRIOR APPLICATION NUMBER: US 60/316,031
;  PRIOR FILING DATE: 2001-08-31
;  PRIOR APPLICATION NUMBER: US 60/328,403
;  PRIOR FILING DATE: 2001-10-12
;  PRIOR APPLICATION NUMBER: JP 2001-157043
;  PRIOR FILING DATE: 2001-05-25
;  PRIOR APPLICATION NUMBER: JP 2001-260681
;  PRIOR FILING DATE: 2001-08-30
;  PRIOR APPLICATION NUMBER: JP 2001-313175
;  PRIOR FILING DATE: 2001-10-10
;  NUMBER OF SEQ ID NOS: 488
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 348
;   LENGTH: 239
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-153-668-348

  Query Match             100.0%;  Score 141;  DB 15;  Length 239;
  Best Local Similarity   100.0%;  Pred. No. 9.8e-14;
  Matches   30;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AVPIAQKSEPHSLSSEALMRRAVSLVTDST 30
              ||||||||||||||||||||||||||||||
Db         56 AVPIAQKSEPHSLSSEALMRRAVSLVTDST 85


RESULT 9
US-09-798-116-2
; Sequence 2, Application US/09798116
; Patent No. US20020110851A1
; GENERAL INFORMATION:
;  APPLICANT: Verhagen, Anne Marie
;  APPLICANT:  Ekert, Paul
;  APPLICANT:  Vaux, David
;  TITLE OF INVENTION: No. US20020110851A1el Polypeptides, Modulatory Agents Therefor
;  FILE REFERENCE: 10338-004US
;  CURRENT APPLICATION NUMBER: US/09/798,116
;  CURRENT FILING DATE:  2001-03-02
;  PRIOR APPLICATION NUMBER: AU PQ5995/00
;  PRIOR FILING DATE: 2000-03-02
;  NUMBER OF SEQ ID NOS: 25
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
;   LENGTH: 237
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-798-116-2

  Query Match             97.9%;  Score 138;  DB 10;  Length 237;
  Best Local Similarity   96.7%;  Pred. No. 2.8e-13;
  Matches   29;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AVPIAQKSEPHSLSSEALMRRAVSLVTDST 30
              |||||||||||||:|||||||||||||||
Db         54 AVPIAQKSEPHSLSNEALMRRAVSLVTDST 83


RESULT 10
US-09-798-116-4
; Sequence 4, Application US/09798116
; Patent No. US20020110851A1
; GENERAL INFORMATION:
;  APPLICANT: Verhagen, Anne Marie
;  APPLICANT:  Ekert, Paul
;  APPLICANT:  Vaux, David
;  TITLE OF INVENTION: No. US20020110851A1el Polypeptides, Modulatory Agents Therefor
;  FILE REFERENCE: 10338-004US
;  CURRENT APPLICATION NUMBER: US/09/798,116
;  CURRENT FILING DATE:  2001-03-02
;  PRIOR APPLICATION NUMBER: AU PQ5995/00
;  PRIOR FILING DATE: 2000-03-02
;  NUMBER OF SEQ ID NOS: 25
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
;   LENGTH: 237
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-798-116-4

  Query Match             97.9%;  Score 138;  DB 10;  Length 237;
  Best Local Similarity   96.7%;  Pred. No. 2.8e-13;
  Matches   29;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AVPIAQKSEPHSLSSEALMRRAVSLVTDST 30
              |||||||||||||:|||||||||||||||
Db         54 AVPIAQKSEPHSLSNEALMRRAVSLVTDST 83


RESULT 11
US-09-798-116-9
; Sequence 9, Application US/09798116
; Patent No. US20020110851A1
; GENERAL INFORMATION:
;  APPLICANT: Verhagen, Anne Marie
;  APPLICANT:  Ekert, Paul
;  APPLICANT:  Vaux, David
;  TITLE OF INVENTION: No. US20020110851A1el Polypeptides, Modulatory Agents Therefor
;  FILE REFERENCE: 10338-004US
;  CURRENT APPLICATION NUMBER: US/09/798,116
;  CURRENT FILING DATE:  2001-03-02
;  PRIOR APPLICATION NUMBER: AU PQ5995/00
;  PRIOR FILING DATE: 2000-03-02
```

```
;  CURRENT FILING DATE:  2001-08-24
;  NUMBER OF SEQ ID NOS: 18
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11
;   LENGTH: 35
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-939-293-11

  Query Match             100.0%;  Score 141;  DB 10;  Length 35;
  Best Local Similarity   100.0%;  Pred. No. 1e-14;
  Matches   30;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AVPIAQKSEPHSLSSEALMRRAVSLVTDST 30
              ||||||||||||||||||||||||||||||
Db          1 AVPIAQKSEPHSLSSEALMRRAVSLVTDST 30


RESULT 3
US-09-939-293-8
; Sequence 8, Application US/09939293
; Patent No. US20020132786A1
; GENERAL INFORMATION:
;  APPLICANT: Alnemri, Emad S.
;  TITLE OF INVENTION: AN IAP PEPTIDE OR POLYPEPTIDE
;  TITLE OF INVENTION:  AND METHODS OF USING THE SAME
;  FILE REFERENCE: 480140.465
;  CURRENT APPLICATION NUMBER: US/09/939,293
;  CURRENT FILING DATE:  2001-08-24
;  NUMBER OF SEQ ID NOS: 18
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8
;   LENGTH: 39
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-939-293-8

  Query Match             100.0%;  Score 141;  DB 10;  Length 39;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-14;
  Matches   30;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AVPIAQKSEPHSLSSEALMRRAVSLVTDST 30
              ||||||||||||||||||||||||||||||
Db          1 AVPIAQKSEPHSLSSEALMRRAVSLVTDST 30


RESULT 4
US-09-939-293-2
; Sequence 2, Application US/09939293
; Patent No. US20020132786A1
; GENERAL INFORMATION:
;  APPLICANT: Alnemri, Emad S.
;  TITLE OF INVENTION: AN IAP PEPTIDE OR POLYPEPTIDE
;  TITLE OF INVENTION:  AND METHODS OF USING THE SAME
;  FILE REFERENCE: 480140.465
;  CURRENT APPLICATION NUMBER: US/09/939,293
;  CURRENT FILING DATE:  2001-08-24
;  NUMBER OF SEQ ID NOS: 18
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 40
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-939-293-2

  Query Match             100.0%;  Score 141;  DB 10;  Length 40;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-14;
  Matches   30;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AVPIAQKSEPHSLSSEALMRRAVSLVTDST 30
              ||||||||||||||||||||||||||||||
Db          1 AVPIAQKSEPHSLSSEALMRRAVSLVTDST 30


RESULT 5
US-09-798-116-7
; Sequence 7, Application US/09798116
; Patent No. US20020110851A1
; GENERAL INFORMATION:
;  APPLICANT: Verhagen, Anne Marie
;  APPLICANT:  Ekert, Paul
;  APPLICANT:  Vaux, David
;  TITLE OF INVENTION: No. US20020110851A1el Polypeptides, Modulatory Agents Therefor
;  FILE REFERENCE: 10338-004US
;  CURRENT APPLICATION NUMBER: US/09/798,116
;  CURRENT FILING DATE:  2001-03-02
;  PRIOR APPLICATION NUMBER: AU PQ5995/00
;  PRIOR FILING DATE: 2000-03-02
;  NUMBER OF SEQ ID NOS: 25
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
;   LENGTH: 202
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-798-116-7

  Query Match             100.0%;  Score 141;  DB 10;  Length 202;
  Best Local Similarity   100.0%;  Pred. No. 8e-14;
  Matches   30;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AVPIAQKSEPHSLSSEALMRRAVSLVTDST 30
              ||||||||||||||||||||||||||||||
Db         19 AVPIAQKSEPHSLSSEALMRRAVSLVTDST 48


RESULT 6
US-09-925-297-591
; Sequence 591, Application US/09925297
; Patent No. US20020081659A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA105
;  CURRENT APPLICATION NUMBER: US/09/925,297
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05989
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 928
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 591
;   LENGTH: 227
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (1)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-297-591

  Query Match             100.0%;  Score 141;  DB 9;  Length 227;
  Best Local Similarity   100.0%;  Pred. No. 9.2e-14;
  Matches   30;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AVPIAQKSEPHSLSSEALMRRAVSLVTDST 30
              ||||||||||||||||||||||||||||||
Db         44 AVPIAQKSEPHSLSSEALMRRAVSLVTDST 73


RESULT 7
US-10-141-618-14
; Sequence 14, Application US/10141618
; Publication No. US20030165887A1
; GENERAL INFORMATION:
;  APPLICANT: Reed, John C.
```

GenCore version 5.1.6


OM protein - protein search, using sw model

Run on: October 2, 2003, 09:36:58 ; Search time 24 Seconds
(without alignments)
197.766 Million cell updates/sec

Title: US-09-939-293A-19_COPY_56_85
Perfect score: 141
Sequence: 1 AVPIAQKSEPHSLSSEALMRRAVSLVTDST 30

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 587654 seqs, 158212981 residues

Total number of hits satisfying chosen parameters: 587654

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 141 | 100.0 | 30 | 10 | US-09-939-293-7 | Sequence 7, Appli |
| 2 | 141 | 100.0 | 35 | 10 | US-09-939-293-11 | Sequence 11, Appl |
| 3 | 141 | 100.0 | 39 | 10 | US-09-939-293-8 | Sequence 8, Appli |
| 4 | 141 | 100.0 | 40 | 10 | US-09-939-293-2 | Sequence 2, Appli |
| 5 | 141 | 100.0 | 202 | 10 | US-09-798-116-7 | Sequence 7, Appli |
| 6 | 141 | 100.0 | 227 | 9 | US-09-925-297-591 | Sequence 591, App |
| 7 | 141 | 100.0 | 239 | 12 | US-10-141-618-14 | Sequence 14, Appl |
| 8 | 141 | 100.0 | 239 | 15 | US-10-153-668-348 | Sequence 348, App |
| 9 | 138 | 97.9 | 237 | 10 | US-09-798-116-2 | Sequence 2, Appli |
| 10 | 138 | 97.9 | 237 | 10 | US-09-798-116-4 | Sequence 4, Appli |
| 11 | 125 | 88.7 | 84 | 10 | US-09-798-116-9 | Sequence 9, Appli |
| 12 | 108 | 76.6 | 177 | 10 | US-09-798-116-8 | Sequence 8, Appli |
| 13 | 105 | 74.5 | 177 | 10 | US-09-798-116-6 | Sequence 6, Appli |
| 14 | 70 | 49.6 | 15 | 14 | US-10-068-569-8 | Sequence 8, Appli |
| 15 | 70 | 49.6 | 15 | 15 | US-10-197-634-8 | Sequence 8, Appli |
| 16 | 63 | 44.7 | 13 | 10 | US-09-798-116-20 | Sequence 20, Appl |
| 17 | 63 | 44.7 | 13 | 10 | US-09-798-116-22 | Sequence 22, Appl |
| 18 | 56 | 39.7 | 73 | 10 | US-09-798-116-10 | Sequence 10, Appl |
| 19 | 49 | 34.8 | 10 | 10 | US-09-965-967-18 | Sequence 18, Appl |
| 20 | 49 | 34.8 | 13 | 10 | US-09-965-967-25 | Sequence 25, Appl |
| 21 | 48.5 | 34.4 | 1144 | 15 | US-10-156-761-7801 | Sequence 7801, Ap |
| 22 | 44 | 31.2 | 1177 | 15 | US-10-193-692-4 | Sequence 4, Appli |
| 23 | 44 | 31.2 | 1186 | 15 | US-10-193-692-2 | Sequence 2, Appli |
| 24 | 44 | 31.2 | 1415 | 9 | US-09-815-242-11036 | Sequence 11036, A |
| 25 | 44 | 31.2 | 1518 | 10 | US-09-801-368-152 | Sequence 152, App |
| 26 | 43.5 | 30.9 | 342 | 14 | US-10-001-857-201 | Sequence 201, App |
| 27 | 43 | 30.5 | 105 | 15 | US-10-078-090-188 | Sequence 188, App |
| 28 | 43 | 30.5 | 159 | 12 | US-09-890-688-110 | Sequence 110, App |
| 29 | 43 | 30.5 | 184 | 9 | US-09-925-299-1546 | Sequence 1546, Ap |
| 30 | 43 | 30.5 | 184 | 11 | US-09-925-299-1546 | Sequence 1546, Ap |
| 31 | 43 | 30.5 | 237 | 9 | US-09-925-301-867 | Sequence 867, App |
| 32 | 43 | 30.5 | 334 | 10 | US-09-953-342-24 | Sequence 24, Appl |
| 33 | 43 | 30.5 | 526 | 12 | US-10-021-425-27 | Sequence 27, Appl |
| 34 | 43 | 30.5 | 570 | 8 | US-08-825-486-2 | Sequence 2, Appli |
| 35 | 43 | 30.5 | 570 | 8 | US-08-870-434-7 | Sequence 7, Appli |
| 36 | 43 | 30.5 | 570 | 10 | US-09-372-044-2 | Sequence 2, Appli |
| 37 | 43 | 30.5 | 570 | 11 | US-09-560-150-7 | Sequence 7, Appli |
| 38 | 43 | 30.5 | 570 | 15 | US-10-067-741-7 | Sequence 7, Appli |
| 39 | 43 | 30.5 | 578 | 15 | US-10-106-698-4636 | Sequence 4636, Ap |
| 40 | 43 | 30.5 | 578 | 15 | US-10-137-418A-3 | Sequence 3, Appli |
| 41 | 43 | 30.5 | 941 | 12 | US-10-032-585-7930 | Sequence 7930, Ap |
| 42 | 43 | 30.5 | 1260 | 15 | US-10-245-802-8 | Sequence 8, Appli |
| 43 | 43 | 30.5 | 1390 | 15 | US-10-137-418A-2 | Sequence 2, Appli |
| 44 | 42.5 | 30.1 | 530 | 15 | US-10-156-761-8391 | Sequence 8391, Ap |
| 45 | 42 | 29.8 | 125 | 10 | US-09-764-877-1372 | Sequence 1372, Ap |

ALIGNMENTS

RESULT 1
US-09-939-293-7
; Sequence 7, Application US/09939293
; Patent No. US20020132786A1
; GENERAL INFORMATION:
; APPLICANT: Alnemri, Emad S.
; TITLE OF INVENTION: AN IAP PEPTIDE OR POLYPEPTIDE
; TITLE OF INVENTION: AND METHODS OF USING THE SAME
; FILE REFERENCE: 480140.465
; CURRENT APPLICATION NUMBER: US/09/939,293
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-939-293-7

Query Match 100.0%; Score 141; DB 10; Length 30;
Best Local Similarity 100.0%; Pred. No. 8.5e-15;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AVPIAQKSEPHSLSSEALMRRAVSLVTDST 30
||||||||||||||||||||||||||||||
Db 1 AVPIAQKSEPHSLSSEALMRRAVSLVTDST 30

RESULT 2
US-09-939-293-11
; Sequence 11, Application US/09939293
; Patent No. US20020132786A1
; GENERAL INFORMATION:
; APPLICANT: Alnemri, Emad S.
; TITLE OF INVENTION: AN IAP PEPTIDE OR POLYPEPTIDE
; TITLE OF INVENTION: AND METHODS OF USING THE SAME
; FILE REFERENCE: 480140.465
; CURRENT APPLICATION NUMBER: US/09/939,293

```
FT   Modified-site   9
FT                   /note= "optional C-terminal protecting group"
XX
PN   WO200230959-A2.
XX
PD   18-APR-2002.
XX
PF   12-OCT-2001; 2001WO-US32121.
XX
PR   13-OCT-2000; 2000US-0687549.
XX
PA   (ABBO ) ABBOTT LAB.
XX
PI   Fesik SW,  Meadows RP,  Betz SP,  Liu Z,  Olejniczak ET,  Sun C;
XX
DR   WPI; 2002-444169/47.
XX
PT   Novel peptide derived from wild-type human second mitochondria derived
PT   activator of caspase protein useful for identifying candidate
PT   substances to kill cancerous cells -
XX
PS   Example 1; Page 15; 26pp; English.
XX
CC   The present sequence is a peptide derived from human second
CC   mitochondria derived activator of caspase (smac), also known as
CC   direct inhibitor of apoptosis binding protein with low pI
CC   (DIABLO), but with the native N-terminal alanine residue (see
CC   ABB76209) acetylated.  Claimed smac-derived peptides (see
CC   ABB76208-19) bind to the Bir2 and Bir3 domain of XIAP, an
CC   inhibitor of apoptosis protein (IAP) family member.  Modification
CC   of the N-terminal alanine destroys all binding affinity for the
CC   protein.  Thus, Kd values for Bir-3 and Bir-2 were each over 1,000
CC   uM for the present peptide, compared with 0.43 +/- 0.06 uM and 6.0
CC   +/- 0.9 uM, respectively, for the corresponding wild-type peptide.
CC   The claimed smac-derived peptides can be used to identify candidate
CC   substances which induce or promote apoptosis in cells.  The assay
CC   involves determination of the ability of candidate compounds to
CC   disrupt the binding interaction between a smac peptide and an IAP
CC   family member.
XX
SQ   Sequence   9 AA;

  Query Match             29.8%;  Score 42;  DB 23;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 9.3e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AVPIAQKSE 9
              |||||||||
Db          1 AVPIAQKSE 9


RESULT 8
ABB76228
ID   ABB76228 standard; Peptide; 10 AA.
XX
AC   ABB76228;
XX
DT   09-AUG-2002  (first entry)
XX
DE   Fluorosceinated smac (DIABLO) derived peptide.
XX
KW   DIABLO; smac; inhibitor of apoptosis protein; IAP; apoptosis;
KW   human; cancer; cytostatic; mutant; mutein.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "N-terminal fluorescein"
XX
PN   WO200230959-A2.
XX
PD   18-APR-2002.
XX
PF   12-OCT-2001; 2001WO-US32121.
XX
PR   13-OCT-2000; 2000US-0687549.
XX
PA   (ABBO ) ABBOTT LAB.
XX
PI   Fesik SW,  Meadows RP,  Betz SP,  Liu Z,  Olejniczak ET,  Sun C;
XX
DR   WPI; 2002-444169/47.
XX
PT   Novel peptide derived from wild-type human second mitochondria derived
PT   activator of caspase protein useful for identifying candidate
PT   substances to kill cancerous cells -
XX
PS   Example 1; Page 14; 26pp; English.
XX
CC   The present sequence corresponds to amino acids 1-9 of human
CC   second mitochondria derived activator of caspase (smac), also known
CC   as direct inhibitor of apoptosis binding protein with low pI
CC   (DIABLO), but is fluorosceinated.  The peptide was used in a
CC   fluorescence polarisation-based competition assay designed to
CC   determine the binding affinity of variant smac peptides (see
CC   ABB76206-27) to the Bir-3 and Bir-2 domains of XIAP, an inhibitor
CC   of apoptosis protein (IAP) family member.  Claimed smac-derived
CC   peptides can be used to identify candidate substances which induce
CC   or promote apoptosis in cells.  The assay involves determination of
CC   the ability of candidate compounds to disrupt the binding
CC   interaction between a smac peptide and an IAP family member.
XX
SQ   Sequence   10 AA;

  Query Match             29.8%;  Score 42;  DB 23;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 2.3;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AVPIAQKSE 9
              |||||||||
Db          1 AVPIAQKSE 9


RESULT 9
ABG72319
ID   ABG72319 standard; Peptide; 29 AA.
XX
AC   ABG72319;
XX
DT   29-JAN-2003  (first entry)
XX
DE   Human pro-apoptotic protein DIABLO peptide sequence #15.
XX
KW   Human; pro-apoptotic protein; DIABLO; cell death; apoptosis;
KW   inhibitor of apoptosis; IAP; cancer; vascular disease; hepatic disease;
KW   autoimmune disease; neurodegenerative disease; tissue damage;
KW   muscular tissue damage; heart attack; hepatic tissue damage;
KW   liver disease; immunogen.
XX
OS   Homo sapiens.
XX
PN   US2002110851-A1.
XX
PD   15-AUG-2002.
XX
PF   02-MAR-2001; 2001US-0798116.
XX
PR   02-MAR-2000; 2000AU-0005995.
XX
PA   (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
XX
PI   Verhagen AM,  Ekert PG,  Vaux DL;
XX
DR   WPI; 2003-074681/07.
```

```
XX
PT   New pro-apoptotic polypeptide, useful for screening for agents which
PT   modulate cell death and for treating conditions associated with cell
PT   death or apoptosis e.g. cancer -
XX
PS   Example 8; Page 4; 50pp; English.
XX
CC   The invention relates to an isolated pro-apoptotic polypeptide,
CC   designated DIABLO, or its biologically active fragment of 8 amino acids
CC   in length. Also included are the polynucleotide encoding DIABLO,
CC   expression vectors, transformed host cells, producing a biologically
CC   active fragment of DIABLO (by contacting an inhibitor of apoptosis (IAP)
CC   with a fragment of the polypeptide, and detecting a reduction in activity
CC   of the IAP), producing a natural or synthetic variant of DIABLO
CC   with cell death activity or which reduces IAP activity, an antigen-
CC   binding molecule that specifically binds to DIABLO or its fragment,
CC   detecting DIABLO in a biological sample (by contacting the sample
CC   with an IAP and detecting the presence of an IAP/DIABLO complex),
CC   modulating the death of a cell (by contacting a cell with an
CC   agent, which modulates the level and/or functional activity of a
CC   polypeptide), a composition for treatment/prophylaxis of a DIABLO related
CC   condition comprising an agent which reduces the level/activity of a
CC   polypeptide or DIABLO. DIABLO, or a nucleic acid encoding DIABLO, is
CC   useful for screening for an agent which modulates cell death. An
CC   antigen-binding molecule is useful for detecting DIABLO in a biological
CC   sample. The agent which modulates the level and/or functional activity of
CC   a polypeptide comprising mature or pro-human DIABLO polypeptide, is
CC   useful for the treatment and/or prophylaxis of a condition associated
CC   with expression or activation of DIABLO, such as cancer, vascular
CC   disease, hepatic disease, autoimmune disease and neurodegenerative
CC   disease, tissue damage or muscular tissue damage associated with heart
CC   attack, or hepatic tissue damage associated with a liver disease.
CC   DIABLO is also useful for treatment and/or prophylaxis of conditions
CC   associated with cell death or apoptosis. The present sequence
CC   represents a partial peptide sequence from human DIABLO, identified
CC   by protein sequencing of a protein (later identified as DIABLO) which
CC   co-precipitates with the human IAP protein MIHA (not defined).
XX
SQ   Sequence   29 AA;

  Query Match             28.4%;  Score 40;  DB 24;  Length 29;
  Best Local Similarity   100.0%;  Pred. No. 21;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         22 AVSLVTDST 30
              |||||||||
Db          1 AVSLVTDST 9


RESULT 10
ABG24798
ID   ABG24798 standard; Protein; 30 AA.
XX
AC   ABG24798;
XX
DT   18-FEB-2002  (first entry)
XX
DE   Novel human diagnostic protein #24789.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US08631.
XX
PR   31-MAR-2000; 2000US-0540217.
PR   23-AUG-2000; 2000US-0649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   N-PSDB; AAS88985.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity -
XX
PS   Claim 20; SEQ ID No 55157; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and
CC   polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC   polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC   and gene mapping, and in recombinant production of (II). The
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. ABG00010-ABG30377 represent novel human
CC   diagnostic amino acid sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   30 AA;

  Query Match             27.7%;  Score 39;  DB 22;  Length 30;
  Best Local Similarity   26.1%;  Pred. No. 32;
  Matches    6;  Conservative    9;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          2 VPIAQKSEPHSLSSEALMRRAVS 24
              :|: |:   |:::  | :|: :|
Db          1 LPVHQQMRMHNVAGRATVRQQIS 23


RESULT 11
ABB76218
ID   ABB76218 standard; Peptide; 9 AA.
XX
AC   ABB76218;
XX
DT   09-AUG-2002  (first entry)
XX
DE   Human smac (DIABLO) derived peptide.
XX
KW   DIABLO; smac; inhibitor of apoptosis protein; IAP; apoptosis;
KW   human; cancer; cytostatic; mutant; mutein.
XX
OS   Homo sapiens;
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 5
FT                   /note= "wild-type Ala substituted by Phe"
FT   Modified-site   9
FT                   /note= "optional C-terminal protecting group"
XX
PN   WO200230959-A2.
XX
PD   18-APR-2002.
XX
```

DR WPI; 2002-304115/34.
XX
PT Novel Smac peptides and polynucleotides encoding the peptides, useful
PT for stimulating apoptosis in neoplastic or tumour cell which
PT overexpresses inhibitor of caspase, and for identifying apoptosis
PT modulating compounds -
XX
PS Example 3; Fig 7; 78pp; English.
XX
CC The invention relates to an isolated Smac peptide or polypeptide (I)
CC and an isolated nucleic acid (II) encoding (I). Also described is a
CC method of identifying a compound that inhibits apoptosis, comprising:
CC (a) separately contacting several cell populations expressing a
CC cytosolic Smac (a Smac isoform that begins with MKSDFYF sequence,
CC replacing the mitochondrial targeting sequence (residues 1-55 of (I)),
CC and residues 56-60 of (I)) and an inhibitor of BID (Bcl2 interacting
CC domain) with a compound to be tested for apoptotic inhibiting activity;
CC (b) incubating the cell populations with a direct stimulus of the cell
CC death pathway; and (c) measuring the specific apoptotic activity of the
CC cell populations, where inhibition of the specific apoptotic activity is
CC indicative that the compound is an inhibitor of apoptosis. (I) and (II)
CC are useful for inducing apoptosis in a cell. The Smac polypeptide and
CC polynucleotide are useful for stimulating apoptosis in a neoplastic or
CC tumour cell which overexpresses an inhibitor of caspase, where the
CC inhibitor inhibits activation or activity of caspase-3, caspase-7 or
CC caspase-9. Preferably, the cell overexpresses at least a portion of IAP.
CC (I) is useful for identifying an inhibitor or enhancer of a caspase-
CC mediated apoptosis which involves contacting a cell transformed or
CC transfected with a vector expressing (I) with a candidate inhibitor or
CC candidate enhancer; and detecting cell viability, where an increase in
CC cell viability indicates the presence of an inhibitor and a decrease in
CC cell viability indicates the presence of an enhancer. Optionally, the
CC method involves detecting the presence of large and small caspase
CC subunits after contacting cell transformed with the vector expressing
CC (I), with the candidate compound. A decrease in processing indicates the
CC presence of an inhibitor and an increase in the processing indicates the
CC presence of an enhancer. Preferably, the large and small subunits of
CC caspase-3, caspase-7 or caspase-9 are detected. (I) is also useful for
CC identifying a compound that inhibits Smac binding to Smac-binding
CC molecule (a portion of IAP e.g. a BIR domain such as BIR1, BIR2 or BIR3,
CC or a full-length IAP). (II) is useful in gene therapy techniques. The
CC present sequence represents the N-terminal amino acid sequence of Smac
CC protein.
XX
SQ Sequence 40 AA;

Query Match 100.0%; Score 33; DB 23; Length 40;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AVPIAQK 7
||||||| 
Db 1 AVPIAQK 7

RESULT 14
ABG72303
ID ABG72303 standard; Protein; 84 AA.
XX
AC ABG72303;
XX
DT 29-JAN-2003 (first entry)
XX
DE Rat partial sequence for pro-apoptotic protein DIABLO.
XX
KW Rat; pro-apoptotic protein; DIABLO; cell death; apoptosis;
KW inhibitor of apoptosis; IAP; cancer; vascular disease; hepatic disease;
KW autoimmune disease; neurodegenerative disease; tissue damage;
KW muscular tissue damage; heart attack; hepatic tissue damage;
KW liver disease; immunogen.
XX
OS Rattus sp.

PN US2002110851-A1.
XX
PD 15-AUG-2002.
XX
PF 02-MAR-2001; 2001US-0798116.
XX
PR 02-MAR-2000; 2000AU-0005995.
XX
PA (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
XX
PI Verhagen AM, Ekert PG, Vaux DL;
XX
DR WPI; 2003-074681/07.
XX
PT New pro-apoptotic polypeptide, useful for screening for agents which
PT modulate cell death and for treating conditions associated with cell
PT death or apoptosis e.g. cancer -
XX
PS Disclosure; Page 35; 50pp; English.
XX
CC The invention relates to an isolated pro-apoptotic polypeptide,
CC designated DIABLO, or its biologically active fragment of 8 amino acids
CC in length. Also included are the polynucleotide encoding DIABLO,
CC expression vectors, transformed host cells, producing a biologically
CC active fragment of DIABLO (by contacting an inhibitor of apoptosis (IAP)
CC with a fragment of the polypeptide, and detecting a reduction in activity
CC of the IAP), producing a natural or synthetic variant of DIABLO
CC with cell death activity or which reduces IAP activity, an antigen-
CC binding molecule that specifically binds to DIABLO or its fragment,
CC detecting DIABLO in a biological sample (by contacting the sample
CC with an IAP and detecting the presence of an IAP/DIABLO complex),
CC modulating the death of a cell (by contacting a cell with an
CC agent, which modulates the level and/or functional activity of a
CC polypeptide), a composition for treatment/prophylaxis of a DIABLO related
CC condition comprising an agent which reduces the level/activity of a
CC polypeptide or DIABLO. DIABLO, or a nucleic acid encoding DIABLO, is
CC useful for screening for an agent which modulates cell death. An
CC antigen-binding molecule is useful for detecting DIABLO in a biological
CC sample. The agent which modulates the level and/or functional activity of
CC a polypeptide comprising mature or pro-human DIABLO polypeptide, is
CC useful for the treatment and/or prophylaxis of a condition associated
CC with expression or activation of DIABLO, such as cancer, vascular
CC disease, hepatic disease, autoimmune disease and neurodegenerative
CC disease, tissue damage or muscular tissue damage associated with heart
CC attack, or hepatic tissue damage associated with a liver disease.
CC DIABLO is also useful for treatment and/or prophylaxis of conditions
CC associated with cell death or apoptosis. The present sequence
CC represents partial rat DIABLO.
XX
SQ Sequence 84 AA;

Query Match 100.0%; Score 33; DB 24; Length 84;
Best Local Similarity 100.0%; Pred. No. 2.9;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AVPIAQK 7
||||||| 
Db 54 AVPIAQK 60

RESULT 15
ABG72302
ID ABG72302 standard; Protein; 202 AA.
XX
AC ABG72302;
XX
DT 29-JAN-2003 (first entry)
XX
DE Human partial sequence for pro-apoptotic protein DIABLO.
XX
KW Human; pro-apoptotic protein; DIABLO; cell death; apoptosis;
KW inhibitor of apoptosis; IAP; cancer; vascular disease; hepatic disease;
KW autoimmune disease; neurodegenerative disease; tissue damage;

OM protein - protein search, using sw model

Run on: October 2, 2003, 09:37:28 ; Search time 35 Seconds
(without alignments)
136.051 Million cell updates/sec

Title: US-09-939-293A-19_COPY_56_85
Perfect score: 141
Sequence: 1 AVPIAQKSEPHSLSSEALMRRAVSLVTDST 30

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 465619

Minimum DB seq length: 0
Maximum DB seq length: 30

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 141 | 100.0 | 30 | 23 | AAU78435 | Inhibitor of apopt |
| 2 | 96 | 68.1 | 20 | 23 | ABB76208 | Human smac (DIABLO |
| 3 | 70 | 49.6 | 15 | 24 | ABP71314 | Human Smac protein |
| 4 | 63 | 44.7 | 13 | 24 | ABG72314 | Human pro-apoptoti |
| 5 | 63 | 44.7 | 13 | 24 | ABG72316 | Human pro-apoptoti |
| 6 | 42 | 29.8 | 9 | 23 | ABB76209 | Human smac (DIABLO |
| 7 | 42 | 29.8 | 9 | 23 | ABB76229 | Human smac (DIABLO |
| 8 | 42 | 29.8 | 10 | 23 | ABB76228 | Fluorosceinated sm |
| 9 | 40 | 28.4 | 29 | 24 | ABG72319 | Human pro-apoptoti |
| 10 | 39 | 27.7 | 30 | 22 | ABG24798 | Novel human diagno |
| 11 | 38 | 27.0 | 9 | 23 | ABB76218 | Human smac (DIABLO |
| 12 | 38 | 27.0 | 9 | 23 | ABB76221 | Human smac (DIABLO |
| 13 | 38 | 27.0 | 9 | 23 | ABB76222 | Human smac (DIABLO |
| 14 | 38 | 27.0 | 9 | 23 | ABB76225 | Human smac (DIABLO |
| 15 | 38 | 27.0 | 9 | 23 | ABB76226 | Human smac (DIABLO |
| 16 | 38 | 27.0 | 9 | 23 | ABB76227 | Human smac (DIABLO |
| 17 | 37 | 26.2 | 8 | 23 | ABB76212 | Human smac (DIABLO |
| 18 | 37 | 26.2 | 9 | 23 | ABB76224 | Human smac (DIABLO |
| 19 | 37 | 26.2 | 25 | 22 | ABG59643 | Human liver peptid |
| 20 | 37 | 26.2 | 25 | 22 | ABB44274 | Peptide #11780 enc |
| 21 | 36 | 25.5 | 9 | 23 | ABB76210 | Human smac (DIABLO |
| 22 | 36 | 25.5 | 9 | 23 | ABB76211 | Human smac (DIABLO |
| 23 | 36 | 25.5 | 9 | 23 | ABB76216 | Human smac (DIABLO |
| 24 | 34 | 24.1 | 9 | 23 | ABB76223 | Human smac (DIABLO |
| 25 | 33.5 | 23.8 | 26 | 18 | AAW25067 | BRCA2 cancer susce |
| 26 | 33.5 | 23.8 | 29 | 22 | ABG55563 | Human liver peptid |
| 27 | 33.5 | 23.8 | 29 | 22 | ABB40307 | Peptide #7813 enco |
| 28 | 33.5 | 23.8 | 29 | 22 | ABB24710 | Protein #6709 enco |
| 29 | 33.5 | 23.8 | 29 | 22 | AAM61105 | Human brain expres |
| 30 | 33.5 | 23.8 | 29 | 22 | AAM73813 | Human bone marrow |
| 31 | 33.5 | 23.8 | 29 | 22 | AAM20109 | Peptide #6543 enco |
| 32 | 33.5 | 23.8 | 29 | 22 | AAM33999 | Peptide #8036 enco |
| 33 | 33.5 | 23.8 | 29 | 23 | ABG43702 | Human peptide enco |
| 34 | 33.5 | 23.8 | 30 | 23 | AAU85044 | Human MAGE-3 segme |
| 35 | 33 | 23.4 | 7 | 23 | ABB76213 | Human smac (DIABLO |
| 36 | 33 | 23.4 | 7 | 23 | AAU78434 | Inhibitor of apopt |
| 37 | 33 | 23.4 | 7 | 23 | AAU78487 | Smac-7 AV peptoid. |
| 38 | 33 | 23.4 | 24 | 20 | AAY45327 | Human secreted pro |
| 39 | 33 | 23.4 | 26 | 22 | AAM83269 | Human immune/haema |
| 40 | 33 | 23.4 | 30 | 14 | AAR42319 | EBV VCA peptide. |
| 41 | 33 | 23.4 | 30 | 16 | AAR74988 | Epstein-Barr virus |
| 42 | 33 | 23.4 | 30 | 20 | AAW99319 | Epstein-Barr virus |
| 43 | 32.5 | 23.0 | 27 | 21 | AAY95833 | Native human LAMP- |
| 44 | 32 | 22.7 | 9 | 23 | ABB76219 | Human smac (DIABLO |
| 45 | 32 | 22.7 | 20 | 22 | ABB36299 | Peptide #3805 enco |

ALIGNMENTS

RESULT 1
AAU78435
ID AAU78435 standard; Peptide; 30 AA.
XX
AC AAU78435;
XX
DT 18-JUN-2002 (first entry)
XX
DE Inhibitor of apoptosis (IAP) protein Smac, mutant Smac-N30.
XX
KW Human; inhibitor of apoptosis; IAP; Smac; apoptosis; BID; BIR1; BIR2;
KW Bcl2 interacting domain; caspase; BIR domain; BIR3; gene therapy;
KW neoplastic cell; mutant; tumour.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200216418-A2.
XX
PD 28-FEB-2002.
XX
PF 24-AUG-2001; 2001WO-US26492.
XX
PR 24-AUG-2000; 2000US-227735P.
XX
PA (UYJE-) UNIV JEFFERSON THOMAS.
XX
PI Alnemri ES;
XX
DR WPI; 2002-304115/34.
XX
PT Novel Smac peptides and polynucleotides encoding the peptides, useful

```
PT   for stimulating apoptosis in neoplastic or tumour cell which
PT   overexpresses inhibitor of caspase, and for identifying apoptosis
PT   modulating compounds -
XX
PS   Example 3; Fig 7; 78pp; English.
XX
CC   The invention relates to an isolated Smac peptide or polypeptide (I)
CC   and an isolated nucleic acid (II) encoding (I). Also described is a
CC   method of identifying a compound that inhibits apoptosis, comprising:
CC   (a) separately contacting several cell populations expressing a
CC   cytosolic Smac (a Smac isoform that begins with MKSDFYF sequence,
CC   replacing the mitochondrial targeting sequence (residues 1-55 of (I)),
CC   and residues 56-60 of (I)) and an inhibitor of BID (Bcl2 interacting
CC   domain) with a compound to be tested for apoptotic inhibiting activity;
CC   (b) incubating the cell populations with a direct stimulus of the cell
CC   death pathway; and (c) measuring the specific apoptotic activity of the
CC   cell populations, where inhibition of the specific apoptotic activity is
CC   indicative that the compound is an inhibitor of apoptosis. (I) and (II)
CC   are useful for inducing apoptosis in a cell. The Smac polypeptide and
CC   polynucleotide are useful for stimulating apoptosis in a neoplastic or
CC   tumour cell which overexpresses an inhibitor of caspase, where the
CC   inhibitor inhibits activation or activity of caspase-3, caspase-7 or
CC   caspase-9. Preferably, the cell overexpresses at least a portion of IAP.
CC   (I) is useful for identifying an inhibitor or enhancer of a caspase-
CC   mediated apoptosis which involves contacting a cell transformed or
CC   transfected with a vector expressing (I) with a candidate inhibitor or
CC   candidate enhancer; and detecting cell viability, where an increase in
CC   cell viability indicates the presence of an inhibitor and a decrease in
CC   cell viability indicates the presence of an enhancer. Optionally, the
CC   method involves detecting the presence of large and small caspase
CC   subunits after contacting cell transformed with the vector expressing
CC   (I), with the candidate compound. A decrease in processing indicates the
CC   presence of an inhibitor and an increase in the processing indicates the
CC   presence of an enhancer. Preferably, the large and small subunits of
CC   caspase-3, caspase-7 or caspase-9 are detected. (I) is also useful for
CC   identifying a compound that inhibits Smac binding to Smac-binding
CC   molecule (a portion of IAP e.g. a BIR domain such as BIR1, BIR2 or BIR3,
CC   or a full-length IAP). (II) is useful in gene therapy techniques. The
CC   present sequence represents the amino acid sequence of Smac mutant
CC   Smac-N30.
XX
SQ   Sequence   30 AA;

  Query Match             100.0%;  Score 141;  DB 23;  Length 30;
  Best Local Similarity   100.0%;  Pred. No. 2.2e-16;
  Matches   30;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AVPIAQKSEPHSLSSEALMRRAVSLVTDST 30
              ||||||||||||||||||||||||||||||
Db          1 AVPIAQKSEPHSLSSEALMRRAVSLVTDST 30


RESULT 2
ABB76208
ID   ABB76208 standard; Peptide; 20 AA.
XX
AC   ABB76208;
XX
DT   09-AUG-2002  (first entry)
XX
DE   Human smac (DIABLO) derived peptide.
XX
KW   DIABLO; smac; inhibitor of apoptosis protein; IAP; apoptosis;
KW   human; cancer; cytostatic.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   20
FT                   /note= "optional C-terminal protecting group"
XX
PN   WO200230959-A2.
XX
PD   18-APR-2002.
XX
PF   12-OCT-2001; 2001WO-US32121.
XX
PR   13-OCT-2000; 2000US-0687549.
XX
PA   (ABBO ) ABBOTT LAB.
XX
PI   Fesik SW,  Meadows RP,  Betz SP,  Liu Z,  Olejniczak ET,  Sun C;
XX
DR   WPI; 2002-444169/47.
XX
PT   Novel peptide derived from wild-type human second mitochondria derived
PT   activator of caspase protein useful for identifying candidate
PT   substances to kill cancerous cells -
XX
PS   Claim 5; Page 7; 26pp; English.
XX
CC   The present sequence is a peptide derived from wild-type human
CC   second mitochondria derived activator of caspase (smac), also known
CC   as direct inhibitor of apoptosis binding protein with low pI
CC   (DIABLO).  The peptide is one of 12 claimed smac (DIABLO)-derived
CC   peptides (see ABB76208-19) which bind to the Bir2 and Bir3 domain
CC   of XIAP, an inhibitor of apoptosis protein (IAP) family member.
CC   Kd values for Bir-3 and Bir-2 are 0.69 +/- 0.05 uM and 6.7 +/- 0.7
CC   uM, respectively, for the present peptide, compared with 0.42 +/-
CC   0.02 uM and 2.3 +/- 0.3 uM, respectively, for full-length smac.
CC   Modification of the N-terminal alanine destroys binding affinity to
CC   XIAP.  For example, N-terminal acetylation of the present peptide,
CC   replacement of the N-terminal alanine with glycine, propionic acid
CC   or isobityric acid all resulted in Kd values for Bir-3 and for Bir-2
CC   of over 1,000 uM.  The claimed peptides can be used to identify
CC   candidate substances which induce or promote apoptosis in cells.
CC   The assay involves determination of the ability of candidate
CC   compounds to disrupt the binding interaction between a smac (DIABLO)
CC   peptide and an IAP family member.
XX
SQ   Sequence   20 AA;

  Query Match             68.1%;  Score 96;  DB 23;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 4.8e-09;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AVPIAQKSEPHSLSSEALMR 20
              ||||||||||||||||||||
Db          1 AVPIAQKSEPHSLSSEALMR 20


RESULT 3
ABP71314
ID   ABP71314 standard; peptide; 15 AA.
XX
AC   ABP71314;
XX
DT   28-APR-2003  (first entry)
XX
DE   Human Smac protein N-terminal fragment.
XX
KW   Omi; HtrA2; serine protease; inhibitor of apoptosis protein; IAP;
KW   caspase; apoptosis; cytostatic; immunosuppressive; neuroprotective;
KW   vasotropic; gene therapy; Smac.
XX
OS   Homo sapiens.
XX
PN   WO2003006680-A2.
XX
PD   23-JAN-2003.
XX
PF   15-JUL-2002; 2002WO-US22658.
XX
PR   13-JUL-2001; 2001US-305378P.
PR   14-DEC-2001; 2001US-340163P.
XX
```

PA (UYJE-) UNIV JEFFERSON THOMAS.
XX
PI Alnemri ES;
XX
DR WPI; 2003-221760/21.
XX
PT New Omi nucleic acids and peptides that bind to an inhibitor of
PT apoptosis proteins, useful for regulating or altering caspase-mediated
PT apoptosis and for treating cancer, tumor, or autoimmune diseases -
XX
PS Example 2; Fig 6; 83pp; English.
XX
CC The invention relates to polynucleotides encoding an Omi (serine
CC protease) peptide or polypeptide. The Omi peptide specifically binds to a
CC portion of an Inhibitor of Apoptosis Protein (IAP). The Omi polypeptide
CC induces caspase-independent apoptosis, or fails to have serine protease
CC activity. The Omi peptides are useful for regulating or altering
CC apoptosis, specifically caspase-mediated apoptosis, and as immunogens for
CC raising antibodies. Enhancers of apoptosis are useful for treating
CC cancers, tumours or for destroying cells that mediate autoimmune
CC diseases. Compositions may also be used for the treatment of diseases
CC associated with inappropriate activation of apoptosis such as
CC neurodegenerative diseases and ischaemic injury. The antibodies can be
CC used in isolating Omi peptides, polypeptides and their variants, in
CC identifying molecules that interact with Omi peptides and polypeptides,
CC and in inhibiting or enhancing the biological activity of Omi peptides
CC and polypeptides. Sequences ABP71310-315 represent fragments of various
CC IAP-binding proteins, used to determine Omi as a IAP-binding protein.
XX
SQ Sequence 15 AA;

Query Match 49.6%; Score 70; DB 24; Length 15;
Best Local Similarity 93.3%; Pred. No. 7.7e-05;
Matches 14; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AVPIAQKSEPHSLSS 15
||||||||||||||:
Db 1 AVPIAQKSEPHSLSN 15

RESULT 4
ABG72314
ID ABG72314 standard; Peptide; 13 AA.
XX
AC ABG72314;
XX
DT 29-JAN-2003 (first entry)
XX
DE Human pro-apoptotic protein DIABLO peptide sequence #10.
XX
KW Human; pro-apoptotic protein; DIABLO; cell death; apoptosis;
KW inhibitor of apoptosis; IAP; cancer; vascular disease; hepatic disease;
KW autoimmune disease; neurodegenerative disease; tissue damage;
KW muscular tissue damage; heart attack; hepatic tissue damage;
KW liver disease; immunogen.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Modified-site 12
FT /label= OTHER
FT /note= "Methione is methionine sulphoxide"
XX
PN US2002110851-A1.
XX
PD 15-AUG-2002.
XX
PF 02-MAR-2001; 2001US-0798116.
XX
PR 02-MAR-2000; 2000AU-0005995.
XX
PA (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
XX
PI Verhagen AM, Ekert PG, Vaux DL;
XX
DR WPI; 2003-074681/07.
XX
PT New pro-apoptotic polypeptide, useful for screening for agents which
PT modulate cell death and for treating conditions associated with cell
PT death or apoptosis e.g. cancer -
XX
PS Example 8; Page 4; 50pp; English.
XX
CC The invention relates to an isolated pro-apoptotic polypeptide,
CC designated DIABLO, or its biologically active fragment of 8 amino acids
CC in length. Also included are the polynucleotide encoding DIABLO,
CC expression vectors, transformed host cells, producing a biologically
CC active fragment of DIABLO (by contacting an inhibitor of apoptosis (IAP)
CC with a fragment of the polypeptide, and detecting a reduction in activity
CC of the IAP), producing a natural or synthetic variant of DIABLO
CC with cell death activity or which reduces IAP activity, an antigen-
CC binding molecule that specifically binds to DIABLO or its fragment,
CC detecting DIABLO in a biological sample (by contacting the sample
CC with an IAP and detecting the presence of an IAP/DIABLO complex),
CC modulating the death of a cell (by contacting a cell with an
CC agent, which modulates the level and/or functional activity of a
CC polypeptide), a composition for treatment/prophylaxis of a DIABLO related
CC condition comprising an agent which reduces the level/activity of a
CC polypeptide or DIABLO. DIABLO, or a nucleic acid encoding DIABLO, is
CC useful for screening for an agent which modulates cell death. An
CC antigen-binding molecule is useful for detecting DIABLO in a biological
CC sample. The agent which modulates the level and/or functional activity of
CC a polypeptide comprising mature or pro-human DIABLO polypeptide, is
CC useful for the treatment and/or prophylaxis of a condition associated
CC with expression or activation of DIABLO, such as cancer, vascular
CC disease, hepatic disease, autoimmune disease and neurodegenerative
CC disease, tissue damage or muscular tissue damage associated with heart
CC attack, or hepatic tissue damage associated with a liver disease.
CC DIABLO is also useful for treatment and/or prophylaxis of conditions
CC associated with cell death or apoptosis. The present sequence
CC represents a partial peptide sequence from human DIABLO, identified
CC by protein sequencing of a protein (later identified as DIABLO) which
CC co-precipitates with the human IAP protein MIHA (not defined).
XX
SQ Sequence 13 AA;

Query Match 44.7%; Score 63; DB 24; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.00096;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 SEPHSLSSEALMR 20
||||||||||||||
Db 1 SEPHSLSSEALMR 13

RESULT 5
ABG72316
ID ABG72316 standard; Peptide; 13 AA.
XX
AC ABG72316;
XX
DT 29-JAN-2003 (first entry)
XX
DE Human pro-apoptotic protein DIABLO peptide sequence #12.
XX
KW Human; pro-apoptotic protein; DIABLO; cell death; apoptosis;
KW inhibitor of apoptosis; IAP; cancer; vascular disease; hepatic disease;
KW autoimmune disease; neurodegenerative disease; tissue damage;
KW muscular tissue damage; heart attack; hepatic tissue damage;
KW liver disease; immunogen.
XX
OS Homo sapiens.
XX
PN US2002110851-A1.
XX
PD 15-AUG-2002.

XX
PF 02-MAR-2001; 2001US-0798116.
XX
PR 02-MAR-2000; 2000AU-0005995.
XX
PA (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
XX
PI Verhagen AM, Ekert PG, Vaux DL;
XX
DR WPI; 2003-074681/07.
XX
PT New pro-apoptotic polypeptide, useful for screening for agents which
PT modulate cell death and for treating conditions associated with cell
PT death or apoptosis e.g. cancer -
XX
PS Example 8; Page 4; 50pp; English.
XX
CC The invention relates to an isolated pro-apoptotic polypeptide,
CC designated DIABLO, or its biologically active fragment of 8 amino acids
CC in length. Also included are the polynucleotide encoding DIABLO,
CC expression vectors, transformed host cells, producing a biologically
CC active fragment of DIABLO (by contacting an inhibitor of apoptosis (IAP)
CC with a fragment of the polypeptide, and detecting a reduction in activity
CC of the IAP), producing a natural or synthetic variant of DIABLO
CC with cell death activity or which reduces IAP activity, an antigen-
CC binding molecule that specifically binds to DIABLO or its fragment,
CC detecting DIABLO in a biological sample (by contacting the sample
CC with an IAP and detecting the presence of an IAP/DIABLO complex),
CC modulating the death of a cell (by contacting a cell with an
CC agent, which modulates the level and/or functional activity of a
CC polypeptide), a composition for treatment/prophylaxis of a DIABLO related
CC condition comprising an agent which reduces the level/activity of a
CC polypeptide or DIABLO. DIABLO, or a nucleic acid encoding DIABLO, is
CC useful for screening for an agent which modulates cell death. An
CC antigen-binding molecule is useful for detecting DIABLO in a biological
CC sample. The agent which modulates the level and/or functional activity of
CC a polypeptide comprising mature or pro-human DIABLO polypeptide, is
CC useful for the treatment and/or prophylaxis of a condition associated
CC with expression or activation of DIABLO, such as cancer, vascular
CC disease, hepatic disease, autoimmune disease and neurodegenerative
CC disease, tissue damage or muscular tissue damage associated with heart
CC attack, or hepatic tissue damage associated with a liver disease.
CC DIABLO is also useful for treatment and/or prophylaxis of conditions
CC associated with cell death or apoptosis. The present sequence
CC represents a partial peptide sequence from human DIABLO, identified
CC by protein sequencing of a protein (later identified as DIABLO) which
CC co-precipitates with the human IAP protein MIHA (not defined).
XX
SQ Sequence 13 AA;

Query Match 44.7%; Score 63; DB 24; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.00096;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 SEPHSLSSEALMR 20
||||||||||||||
Db 1 SEPHSLSSEALMR 13

RESULT 6
ABB76209
ID ABB76209 standard; Peptide; 9 AA.
XX
AC ABB76209;
XX
DT 09-AUG-2002 (first entry)
XX
DE Human smac (DIABLO) derived peptide.
XX
KW DIABLO; smac; inhibitor of apoptosis protein; IAP; apoptosis;
KW human; cancer; cytostatic.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Modified-site 9
FT /note= "optional C-terminal protecting group"
XX
PN WO200230959-A2.
XX
PD 18-APR-2002.
XX
PF 12-OCT-2001; 2001WO-US32121.
XX
PR 13-OCT-2000; 2000US-0687549.
XX
PA (ABBO ) ABBOTT LAB.
XX
PI Fesik SW, Meadows RP, Betz SP, Liu Z, Olejniczak ET, Sun C;
XX
DR WPI; 2002-444169/47.
XX
PT Novel peptide derived from wild-type human second mitochondria derived
PT activator of caspase protein useful for identifying candidate
PT substances to kill cancerous cells -
XX
PS Claim 5; Page 7; 26pp; English.
XX
CC The present sequence is a peptide derived from wild-type human
CC second mitochondria derived activator of caspase (smac), also known
CC as direct inhibitor of apoptosis binding protein with low pI
CC (DIABLO). The peptide is one of 12 claimed smac (DIABLO)-derived
CC peptides (see ABB76208-19) which bind to the Bir2 and Bir3 domain
CC of XIAP, an inhibitor of apoptosis protein (IAP) family member.
CC Kd values for Bir-3 and Bir-2 are 0.43 +/- 0.06 uM and 6.0 +/- 0.9
CC uM, respectively, for the present peptide, compared with 0.42 +/-
CC 0.02 uM and 2.3 +/- 0.3 uM, respectively, for full-length smac.
CC Modification of the N-terminal alanine destroys binding affinity to
CC XIAP, and mutation of the valine, proline or isoleucine also causes
CC some loss of binding. Amino acids C-terminal to the isoleucine are
CC not important for binding. The claimed peptides can be used to
CC identify candidate substances which induce or promote apoptosis in
CC cells. The assay involves determination of the ability of
CC candidate compounds to disrupt the binding interaction between a
CC smac (DIABLO) peptide and an IAP family member.
XX
SQ Sequence 9 AA;

Query Match 29.8%; Score 42; DB 23; Length 9;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AVPIAQKSE 9
|||||||||
Db 1 AVPIAQKSE 9

RESULT 7
ABB76229
ID ABB76229 standard; Peptide; 9 AA.
XX
AC ABB76229;
XX
DT 09-AUG-2002 (first entry)
XX
DE Human smac (DIABLO) derived peptide.
XX
KW DIABLO; smac; inhibitor of apoptosis protein; IAP; apoptosis;
KW human; cancer; cytostatic; mutant; mutein.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 1
FT /note= "N-terminal acetyl"

```
FT   Modified-site   9
FT                   /note= "optional C-terminal protecting group"
XX
PN   WO200230959-A2.
XX
PD   18-APR-2002.
XX
PF   12-OCT-2001; 2001WO-US32121.
XX
PR   13-OCT-2000; 2000US-0687549.
XX
PA   (ABBO ) ABBOTT LAB.
XX
PI   Fesik SW,  Meadows RP,  Betz SP,  Liu Z,  Olejniczak ET,  Sun C;
XX
DR   WPI; 2002-444169/47.
XX
PT   Novel peptide derived from wild-type human second mitochondria derived
PT   activator of caspase protein useful for identifying candidate
PT   substances to kill cancerous cells -
XX
PS   Example 1; Page 15; 26pp; English.
XX
CC   The present sequence is a peptide derived from human second
CC   mitochondria derived activator of caspase (smac), also known as
CC   direct inhibitor of apoptosis binding protein with low pI
CC   (DIABLO), but with the native N-terminal alanine residue (see
CC   ABB76209) acetylated.  Claimed smac-derived peptides (see
CC   ABB76208-19) bind to the Bir2 and Bir3 domain of XIAP, an
CC   inhibitor of apoptosis protein (IAP) family member.  Modification
CC   of the N-terminal alanine destroys all binding affinity for the
CC   protein.  Thus, Kd values for Bir-3 and Bir-2 were each over 1,000
CC   uM for the present peptide, compared with 0.43 +/- 0.06 uM and 6.0
CC   +/- 0.9 uM, respectively, for the corresponding wild-type peptide.
CC   The claimed smac-derived peptides can be used to identify candidate
CC   substances which induce or promote apoptosis in cells.  The assay
CC   involves determination of the ability of candidate compounds to
CC   disrupt the binding interaction between a smac peptide and an IAP
CC   family member.
XX
SQ   Sequence   9 AA;

  Query Match             29.8%;  Score 42;  DB 23;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 9.3e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AVPIAQKSE 9
              |||||||||
Db          1 AVPIAQKSE 9


RESULT 8
ABB76228
ID   ABB76228 standard; Peptide; 10 AA.
XX
AC   ABB76228;
XX
DT   09-AUG-2002  (first entry)
XX
DE   Fluorosceinated smac (DIABLO) derived peptide.
XX
KW   DIABLO; smac; inhibitor of apoptosis protein; IAP; apoptosis;
KW   human; cancer; cytostatic; mutant; mutein.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT   .               /note= "N-terminal fluorescein"
XX
PN   WO200230959-A2.
XX
PD   18-APR-2002.
XX
PF   12-OCT-2001; 2001WO-US32121.
XX
PR   13-OCT-2000; 2000US-0687549.
XX
PA   (ABBO ) ABBOTT LAB.
XX
PI   Fesik SW,  Meadows RP,  Betz SP,  Liu Z,  Olejniczak ET,  Sun C;
XX
DR   WPI; 2002-444169/47.
XX
PT   Novel peptide derived from wild-type human second mitochondria derived
PT   activator of caspase protein useful for identifying candidate
PT   substances to kill cancerous cells -
XX
PS   Example 1; Page 14; 26pp; English.
XX
CC   The present sequence corresponds to amino acids 1-9 of human
CC   second mitochondria derived activator of caspase (smac), also known
CC   as direct inhibitor of apoptosis binding protein with low pI
CC   (DIABLO), but is fluorosceinated.  The peptide was used in a
CC   fluorescence polarisation-based competition assay designed to
CC   determine the binding affinity of variant smac peptides (see
CC   ABB76206-27) to the Bir-3 and Bir-2 domains of XIAP, an inhibitor
CC   of apoptosis protein (IAP) family member.  Claimed smac-derived
CC   peptides can be used to identify candidate substances which induce
CC   or promote apoptosis in cells.  The assay involves determination of
CC   the ability of candidate compounds to disrupt the binding
CC   interaction between a smac peptide and an IAP family member.
XX
SQ   Sequence   10 AA;

  Query Match             29.8%;  Score 42;  DB 23;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 2.3;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AVPIAQKSE 9
              |||||||||
Db          1 AVPIAQKSE 9


RESULT 9
ABG72319
ID   ABG72319 standard; Peptide; 29 AA.
XX
AC   ABG72319;
XX
DT   29-JAN-2003  (first entry)
XX
DE   Human pro-apoptotic protein DIABLO peptide sequence #15.
XX
KW   Human; pro-apoptotic protein; DIABLO; cell death; apoptosis;
KW   inhibitor of apoptosis; IAP; cancer; vascular disease; hepatic disease;
KW   autoimmune disease; neurodegenerative disease; tissue damage;
KW   muscular tissue damage; heart attack; hepatic tissue damage;
KW   liver disease; immunogen.
XX
OS   Homo sapiens.
XX
PN   US2002110851-A1.
XX
PD   15-AUG-2002.
XX
PF   02-MAR-2001; 2001US-0798116.
XX
PR   02-MAR-2000; 2000AU-0005995.
XX
PA   (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
XX
PI   Verhagen AM,  Ekert PG,  Vaux DL;
XX
DR   WPI; 2003-074681/07.
```

```
XX
PT   New pro-apoptotic polypeptide, useful for screening for agents which
PT   modulate cell death and for treating conditions associated with cell
PT   death or apoptosis e.g. cancer -
XX
PS   Example 8; Page 4; 50pp; English.
XX
CC   The invention relates to an isolated pro-apoptotic polypeptide,
CC   designated DIABLO, or its biologically active fragment of 8 amino acids
CC   in length. Also included are the polynucleotide encoding DIABLO,
CC   expression vectors, transformed host cells, producing a biologically
CC   active fragment of DIABLO (by contacting an inhibitor of apoptosis (IAP)
CC   with a fragment of the polypeptide, and detecting a reduction in activity
CC   of the IAP), producing a natural or synthetic variant of DIABLO
CC   with cell death activity or which reduces IAP activity, an antigen-
CC   binding molecule that specifically binds to DIABLO or its fragment,
CC   detecting DIABLO in a biological sample (by contacting the sample
CC   with an IAP and detecting the presence of an IAP/DIABLO complex),
CC   modulating the death of a cell (by contacting a cell with an
CC   agent, which modulates the level and/or functional activity of a
CC   polypeptide), a composition for treatment/prophylaxis of a DIABLO related
CC   condition comprising an agent which reduces the level/activity of a
CC   polypeptide or DIABLO. DIABLO, or a nucleic acid encoding DIABLO, is
CC   useful for screening for an agent which modulates cell death. An
CC   antigen-binding molecule is useful for detecting DIABLO in a biological
CC   sample. The agent which modulates the level and/or functional activity of
CC   a polypeptide comprising mature or pro-human DIABLO polypeptide, is
CC   useful for the treatment and/or prophylaxis of a condition associated
CC   with expression or activation of DIABLO, such as cancer, vascular
CC   disease, hepatic disease, autoimmune disease and neurodegenerative
CC   disease, tissue damage or muscular tissue damage associated with heart
CC   attack, or hepatic tissue damage associated with a liver disease.
CC   DIABLO is also useful for treatment and/or prophylaxis of conditions
CC   associated with cell death or apoptosis. The present sequence
CC   represents a partial peptide sequence from human DIABLO, identified
CC   by protein sequencing of a protein (later identified as DIABLO) which
CC   co-precipitates with the human IAP protein MIHA (not defined).
XX
SQ   Sequence   29 AA;

  Query Match             28.4%;  Score 40;  DB 24;  Length 29;
  Best Local Similarity   100.0%;  Pred. No. 21;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         22 AVSLVTDST 30
              |||||||||
Db          1 AVSLVTDST 9


RESULT 10
ABG24798
ID   ABG24798 standard; Protein; 30 AA.
XX
AC   ABG24798;
XX
DT   18-FEB-2002  (first entry)
XX
DE   Novel human diagnostic protein #24789.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US08631.
XX
PR   31-MAR-2000; 2000US-0540217.
PR   23-AUG-2000; 2000US-0649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   N-PSDB; AAS88985.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity -
XX
PS   Claim 20; SEQ ID No 55157; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and
CC   polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC   polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC   and gene mapping, and in recombinant production of (II). The
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. ABG00010-ABG30377 represent novel human
CC   diagnostic amino acid sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   30 AA;

  Query Match             27.7%;  Score 39;  DB 22;  Length 30;
  Best Local Similarity   26.1%;  Pred. No. 32;
  Matches    6;  Conservative    9;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          2 VPIAQKSEPHSLSSEALMRRAVS 24
              :|: |:   |::  | :|: :|
Db          1 LPVHQQMRMHNVAGRATVRQQIS 23


RESULT 11
ABB76218
ID   ABB76218 standard; Peptide; 9 AA.
XX
AC   ABB76218;
XX
DT   09-AUG-2002  (first entry)
XX
DE   Human smac (DIABLO) derived peptide.
XX
KW   DIABLO; smac; inhibitor of apoptosis protein; IAP; apoptosis;
KW   human; cancer; cytostatic; mutant; mutein.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 5
FT                   /note= "wild-type Ala substituted by Phe"
FT   Modified-site   9
FT                   /note= "optional C-terminal protecting group"
XX
PN   WO200230959-A2.
XX
PD   18-APR-2002.
XX
```

```
PF   12-OCT-2001; 2001WO-US32121.
XX
PR   13-OCT-2000; 2000US-0687549.
XX
PA   (ABBO ) ABBOTT LAB.
XX
PI   Fesik SW,  Meadows RP,  Betz SP,  Liu Z,  Olejniczak ET,  Sun C;
XX
DR   WPI; 2002-444169/47.
XX
PT   Novel peptide derived from wild-type human second mitochondria derived
PT   activator of caspase protein useful for identifying candidate
PT   substances to kill cancerous cells -
XX
PS   Claim 5; Page 7; 26pp; English.
XX
CC   The present sequence is a peptide derived from wild-type human
CC   second mitochondria derived activator of caspase (smac), also known
CC   as direct inhibitor of apoptosis binding protein with low pI
CC   (DIABLO).  The peptide is one of 12 claimed smac (DIABLO)-derived
CC   peptides (see ABB76208-19) which bind to the Bir2 and Bir3 domain
CC   of XIAP, an inhibitor of apoptosis protein (IAP) family member.
CC   Kd values for Bir-3 and Bir-2 are 0.5 +/- 0.1 uM and 2.5 +/- 5.0
CC   uM, respectively, for the present peptide, compared with 0.42 +/-
CC   0.02 uM and 2.3 +/- 0.3 uM, respectively, for full-length smac.
CC   Amino acids C-terminal to the isoleucine of smac (DIABLO) are not
CC   important for binding to XIAP.  The claimed peptides can be used
CC   to identify candidate substances which induce or promote apoptosis
CC   in cells.  The assay involves determination of the ability of
CC   candidate compounds to disrupt the binding interaction between a
CC   smac (DIABLO) peptide and an IAP family member.
XX
SQ   Sequence   9 AA;

  Query Match             27.0%;  Score 38;  DB 23;  Length 9;
  Best Local Similarity   88.9%;  Pred. No. 9.3e+05;
  Matches    8;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 AVPIAQKSE 9
              ||| ||||| 
Db          1 AVPFAQKSE 9


RESULT 12
ABB76221
ID   ABB76221 standard; Peptide; 9 AA.
XX
AC   ABB76221;
XX
DT   09-AUG-2002  (first entry)
XX
DE   Human smac (DIABLO) derived peptide.
XX
KW   DIABLO; smac; inhibitor of apoptosis protein; IAP; apoptosis;
KW   human; cancer; cytostatic; mutant; mutein.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1
FT                   /note= "wild-type Ala substituted by Gly"
FT   Modified-site   9
FT                   /note= "optional C-terminal protecting group"
XX
PN   WO200230959-A2.
XX
PD   18-APR-2002.
XX
PF   12-OCT-2001; 2001WO-US32121.
XX
PR   13-OCT-2000; 2000US-0687549.
XX
PA   (ABBO ) ABBOTT LAB.
XX
PI   Fesik SW,  Meadows RP,  Betz SP,  Liu Z,  Olejniczak ET,  Sun C;
XX
DR   WPI; 2002-444169/47.
XX
PT   Novel peptide derived from wild-type human second mitochondria derived
PT   activator of caspase protein useful for identifying candidate
PT   substances to kill cancerous cells -
XX
PS   Example 1; Page 15; 26pp; English.
XX
CC   The present sequence is a peptide derived from human second
CC   mitochondria derived activator of caspase (smac), also known as
CC   direct inhibitor of apoptosis binding protein with low pI
CC   (DIABLO), but has the native N-terminal alanine residue (see
CC   ABB76209) replaced by glycine.  Claimed smac-derived peptides
CC   (see ABB76208-19) bind to the Bir2 and Bir3 domain of XIAP, an
CC   inhibitor of apoptosis protein (IAP) family member.  Modification
CC   of the N-terminal alanine destroys all binding affinity for the
CC   protein.  Thus, Kd values for Bir-3 and Bir-2 were each over 1,000
CC   uM for the present peptide, compared with 0.43 +/- 0.06 uM and 6.0
CC   +/- 0.9 uM, respectively, for the corresponding N-terminal Ala
CC   peptide.  The claimed smac-derived peptides can be used to identify
CC   candidate substances which induce or promote apoptosis in cells.
CC   The assay involves determination of the ability of candidate
CC   compounds to disrupt the binding interaction between a smac
CC   peptide and an IAP family member.
XX
SQ   Sequence   9 AA;

  Query Match             27.0%;  Score 38;  DB 23;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 9.3e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 VPIAQKSE 9
              ||||||||
Db          2 VPIAQKSE 9


RESULT 13
ABB76222
ID   ABB76222 standard; Peptide; 9 AA.
XX
AC   ABB76222;
XX
DT   09-AUG-2002  (first entry)
XX
DE   Human smac (DIABLO) derived peptide.
XX
KW   DIABLO; smac; inhibitor of apoptosis protein; IAP; apoptosis;
KW   human; cancer; cytostatic; mutant; mutein.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 2
FT                   /note= "wild-type Val substituted by Ala"
FT   Modified-site   9
FT                   /note= "optional C-terminal protecting group"
XX
PN   WO200230959-A2.
XX
PD   18-APR-2002.
XX
PF   12-OCT-2001; 2001WO-US32121.
XX
PR   13-OCT-2000; 2000US-0687549.
XX
PA   (ABBO ) ABBOTT LAB.
XX
PI   Fesik SW,  Meadows RP,  Betz SP,  Liu Z,  Olejniczak ET,  Sun C;
```

```
XX
DR   WPI; 2002-444169/47.
XX
PT   Novel peptide derived from wild-type human second mitochondria derived
PT   activator of caspase protein useful for identifying candidate
PT   substances to kill cancerous cells -
XX
PS   Example 1; Page 15; 26pp; English.
XX
CC   The present sequence is a peptide derived from human second
CC   mitochondria derived activator of caspase (smac), also known as
CC   direct inhibitor of apoptosis binding protein with low pI
CC   (DIABLO), but has the native valine residue (see ABB76209) replaced
CC   by alanine.  Claimed smac-derived peptides (see ABB76208-19) bind
CC   to the Bir2 and Bir3 domain of XIAP, an inhibitor of apoptosis
CC   protein (IAP) family member.  Mutation of the valine of the peptide
CC   causes some loss of binding to the protein.  Thus, Kd values for
CC   Bir-3 and Bir-2 were 12 +/- 2 uM and 56 +/- 5 uM, respectively,
CC   uM for the present peptide, compared with 0.43 +/- 0.06 uM and 6.0
CC   +/- 0.9 uM, respectively, for the corresponding wild-type peptide.
CC   The claimed smac-derived peptides can be used to identify candidate
CC   substances which induce or promote apoptosis in cells.  The assay
CC   involves determination of the ability of candidate compounds to
CC   disrupt the binding interaction between a smac peptide and an IAP
CC   family member.
XX
SQ   Sequence   9 AA;

  Query Match             27.0%;  Score 38;  DB 23;  Length 9;
  Best Local Similarity   88.9%;  Pred. No. 9.3e+05;
  Matches    8;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 AVPIAQKSE 9
              | |||||||
Db          1 AAPIAQKSE 9


RESULT 14
ABB76225
ID   ABB76225 standard; Peptide; 9 AA.
XX
AC   ABB76225;
XX
DT   09-AUG-2002  (first entry)
XX
DE   Human smac (DIABLO) derived peptide.
XX
KW   DIABLO; smac; inhibitor of apoptosis protein; IAP; apoptosis;
KW   human; cancer; cytostatic; mutant; mutein.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 5
FT                   /note= "wild-type Ala substituted by Gly"
FT   Modified-site   9
FT                   /note= "optional C-terminal protecting group"
XX
PN   WO200230959-A2.
XX
PD   18-APR-2002.
XX
PF   12-OCT-2001; 2001WO-US32121.
XX
PR   13-OCT-2000; 2000US-0687549.
XX
PA   (ABBO ) ABBOTT LAB.
XX
PI   Fesik SW,  Meadows RP,  Betz SP,  Liu Z,  Olejniczak ET,  Sun C;
XX
DR   WPI; 2002-444169/47.
XX
PT   Novel peptide derived from wild-type human second mitochondria derived
PT   activator of caspase protein useful for identifying candidate
PT   substances to kill cancerous cells -
XX
PS   Example 1; Page 15; 26pp; English.
XX
CC   The present sequence is a peptide derived from human second
CC   mitochondria derived activator of caspase (smac), also known as
CC   direct inhibitor of apoptosis binding protein with low pI
CC   (DIABLO), but has the native Ala residue (see ABB76209) replaced
CC   by glycine.  Claimed smac-derived peptides (see ABB76208-19) bind
CC   to the Bir2 and Bir3 domain of XIAP, an inhibitor of apoptosis
CC   protein (IAP) family member.  Mutation of the amino acids
CC   C-terminal to the isoleucine residue of the wild-type peptide
CC   caused little loss of binding to the protein.  Thus, Kd values for
CC   Bir-3 and Bir-2 were 1.2 +/- 0.4 uM and 10 +/- 2 uM, respectively,
CC   uM for the present peptide, compared with 0.43 +/- 0.06 uM and 6.0
CC   +/- 0.9 uM, respectively, for the corresponding wild-type peptide.
CC   The claimed smac-derived peptides can be used to identify candidate
CC   substances which induce or promote apoptosis in cells.  The assay
CC   involves determination of the ability of candidate compounds to
CC   disrupt the binding interaction between a smac peptide and an IAP
CC   family member.
XX
SQ   Sequence   9 AA;

  Query Match             27.0%;  Score 38;  DB 23;  Length 9;
  Best Local Similarity   88.9%;  Pred. No. 9.3e+05;
  Matches    8;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 AVPIAQKSE 9
              |||| ||||
Db          1 AVPIGQKSE 9


RESULT 15
ABB76226
ID   ABB76226 standard; Peptide; 9 AA.
XX
AC   ABB76226;
XX
DT   09-AUG-2002  (first entry)
XX
DE   Human smac (DIABLO) derived peptide.
XX
KW   DIABLO; smac; inhibitor of apoptosis protein; IAP; apoptosis;
KW   human; cancer; cytostatic; mutant; mutein.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "n-propionic acid"
FT   Modified-site   9
FT                   /note= "optional C-terminal protecting group"
XX
PN   WO200230959-A2.
XX
PD   18-APR-2002.
XX
PF   12-OCT-2001; 2001WO-US32121.
XX
PR   13-OCT-2000; 2000US-0687549.
XX
PA   (ABBO ) ABBOTT LAB.
XX
PI   Fesik SW,  Meadows RP,  Betz SP,  Liu Z,  Olejniczak ET,  Sun C;
XX
DR   WPI; 2002-444169/47.
XX
PT   Novel peptide derived from wild-type human second mitochondria derived
PT   activator of caspase protein useful for identifying candidate
```

PT substances to kill cancerous cells -
XX
PS Example 1; Page 15; 26pp; English.
XX
CC The present sequence is a peptide derived from human second
CC mitochondria derived activator of caspase (smac), also known as
CC direct inhibitor of apoptosis binding protein with low pI
CC (DIABLO), but has the native N-terminal alanine residue (see
CC ABB76209) replaced by propionic acid. Claimed smac-derived peptides
CC (see ABB76208-19) bind to the Bir2 and Bir3 domain of XIAP, an
CC inhibitor of apoptosis protein (IAP) family member. Modification
CC of the N-terminal alanine destroys all binding affinity for the
CC protein. Thus, Kd values for Bir-3 and Bir-2 were each over 1,000
CC uM for the present peptide, compared with 0.43 +/- 0.06 uM and 6.0
CC +/- 0.9 uM, respectively, for the corresponding N-terminal Ala
CC peptide. The claimed smac-derived peptides can be used to identify
CC candidate substances which induce or promote apoptosis in cells.
CC The assay involves determination of the ability of candidate
CC compounds to disrupt the binding interaction between a smac
CC peptide and an IAP family member.
XX
SQ Sequence 9 AA;

Query Match 27.0%; Score 38; DB 23; Length 9;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          2 VPIAQKSE 9
              ||||||||
Db          2 VPIAQKSE 9
```

Search completed: October 2, 2003, 09:41:55
Job time : 36 secs